#Editor

Comments to the Author:

There seem to be missing corrections asked by referee 2 from version 4 of the manuscript onwards:

These points seem not to be addressed neither in gmd-2019-186-author_response-version2.pdf nor in gmd-2019-186-manuscript-version[5-7].pdf

Maybe they were not included in the uploaded version? Or am I missing something? Until 33) the points are addressed in gmd-2019-186-manuscript-version5.pdf.

There may have been some misunderstanding. We have answered the points from 33) in October 2020, but we have thought that the resulting discussions should not lead to corrections in the manuscript. In the following, we include the answers we gave in gmd-2019-186-author_response-version2.pdf in italic. Further explanations or new answers are given in bold.

The following points seem to be still missing:

34) Line 259: "standard deviation coefficient". Please clarify. Is it really a coefficient? And since this is an error covariance matrix, should the diagonal elements be error variance not error standard deviation?

in gmd-2019-186-author_response-version2.pdf p54: *Indeed, the diagonal elements are variances, as already explained in Section 3.4: "The variances are specified by the user through standard deviation coefficient (Table 1), which can be a fixed value ("fx") or a percentage ("pc") to define the diagonal standard deviation matrix ∑."*

We have changed the sentence: "The variances are specified by the user through the specification of the values for the corresponding standard deviation (i.e. the diagonal matrix of standard deviations ∑, Table 1) which can be made in terms of fixed values ("fx") or percentages ("pc")."

35) Line 260-262: Very important statement. But please elaborate or rephrase. What is standard deviation of the uncertainty?

The different ways of building the error covariance matrix are detailed in gmd-2019-186-author_response-version2.pdf p10-11, from line 308 to 344. The words "standard deviation of the uncertainty" have been removed and this comment does not apply anymore.We have changed the sentence: "For correction types "mult" and "scale", as well as for correction type "add" with a fixed value, the value is directly used as the uncertainty in the corresponding components of the control vector."

36) Line 266: "variances". Are these error variances?

in gmd-2019-186-author_response-version2.pdf p54: *Yes, they are.*

37) Line 270: "error correlation between fluxes of CO and NOx, are not coded yet". Please elaborate on its potential effect on your estimation?

in gmd-2019-186-author_response-version2.pdf p54: *We cannot quantify this potential effect at this stage and to our knowledge, there is no study about this in the literature.*

When handling CO and NO_x emissions from anthropogenic combustion, depending on whether the major source of uncertainty in the CO and NO_x emissions is connected to the level of corresponding activity, or to the emission factors corresponding to the conversion of activity level into emission estimates, such correlations could be high or low, and even negative. Such correlations generate some transfer of information from atmospheric data for one species to the emission of the other species, high positive correlations enhancing the overall constraint on the emissions from a given set of observations. However, there is currently no strong consensus regarding the levels of correlations and it clearly depends on the specific study cases (cities, regions, countries etc.) For other types of emitting processes, one can hardly find some correlations between uncertainties in NO_x and CO fluxes.

We have added a sentence in the text: "Such correlations increase the observation constraint on the emissions in the inversion process by transferring information from one species to the other. The level (and sometimes the sign) and thus the impact on the inversion of such correlations highly depends on the study cases, and are often debated due to the lack of precise characterization of the uncertainties in inventories of anthropogenic emissions of GHG and pollutants [Super et al. 2020]."

38) Line 296: How about calling this "Observation Operators"?

in gmd-2019-186-author_response-version2.pdf p55: *We do not agree. The section title remains "Equivalents of the observations".*

We would like to recall that there is only one observation operator, which includes the CTM but also the various steps (e.g., interpolation, averaging, application of the averaging kernels) to compute the equivalents of the observations. The section title remains: 'Equivalents of the observations".

40) Section 3.4. I think this is very relevant. Please elaborate Figure 3. In its current form, it is not clear what this Figure represents and how we can use it to interpret results. I think coding of these operators is a vital step in the assimilation and should be given more emphasis. Are these utilities also available? How good are the adjoints of these operators? Are there tests to diagnose their accuracy? The reviewer asked for clarifications on figure 3 but it was removed altogether, and further clarifications seem to be missing. Were those considered in an intermediate version?

in gmd-2019-186-author_response-version2.pdf p44: *Figure 3 has been removed, as the Section 3.4 already well describes the calculations of the equivalents of the observations.*

in gmd-2019-186-author_response-version2.pdf p55: *We have added a subsection describing the different tests of our system: "3.3. Accuracy of tangent-linear and adjoint codes".*

The Section 3.3 is recalled here: "Different procedures have been implemented to test the accuracy of the TL and adjoint codes. To test the linearity of the TL, we compute a Taylor diagnostic. It consists in computing the TL at x_0 for given increments Δx, $dHx_0(\Delta x)$, then the TL at x_0 for $\lambda \times \Delta x$ with λ an arbitrary small number, $dHx_0(\lambda \Delta x)$. Theoretically, if the TL is well coded, $\lambda dHx_0(\Delta x) = dHx_0(\lambda \Delta x)$ by definition. In practice, the difference must be lower than 10 times the precision of the machine on which it is run.

The adjoint code is also tested, by verifying that $\langle H.\Delta x, H.\Delta x \rangle = \langle \Delta x, H^T H.\Delta x \rangle$ where H^T stands for the adjoint at x. What is actually computed is the ratio of the difference between the two scalar products to the second one and the accuracy of the computation. The difference should be a few times the precision of the machine on which it is run."

42) Line 314-318: Please highlight in your notations if these are scalars or vectors. And please add corresponding dimensions. What is the difference between small (c_m(o)) and big C_m(o. What is x_a?

Matrices, vectors and scalars are still indicated with the same typeface in v7.

All the components of the following equations are matrices or vectors, described with standard mathematical notations i.e. using bold capital letters for matrices and bold lowercase letters for vectors. We have changed the notations: "Two types of formula, depending on the satellite observations used, have been detailed in PYVAR-CHIMERE for the use of AKs**:** $\boldsymbol{c_m = AK.c_{m(o)}}$ (Eq. 8) **or** $\boldsymbol{c_m = x_a + AK(c_{m(o)} - x_a)}$ (Eq. 9) where $\mathbf{c_m}$ is the modeled column, **AK contains the averaging kernels that can be provided in the form of vector (e.g., OMI product) or matrix (e.g., MOPITT product),** $\boldsymbol{x_a}$ **is the prior state vector** (provided together with the AKs when relevant) and $\mathbf{c_{m(o)}}$ is the vertical distribution of the original model partial columns interpolated to the pressure grid of the AKs.**"**

43) Line 328-334: This is also informative. Is there a reference for parallelization approach in PYVAR and CHIMERE? How does it scale with more CPUs? 4 hours seem to be a long time, isn't it? Please elaborate and compare with other systems.

in gmd-2019-186-author_response-version2.pdf p55:*The parallelization approach for CHIMERE is described in the Section 2.2 of Menut et al. [2013].*

We have added the reference in the text: "As described in Menut et al. [2013] for CHIMERE, the model parallelization results from a Cartesian division of the main geographical domain into several sub-domains, each one being processed by a worker process."

in gmd-2019-186-author_response-version2.pdf p55:*The optimal number of CPUs for the parallelization of the transport scheme depends on the size of the tiles (for the Van Leer scheme, they must be at least 6 grid cell large because of the upwind and downwind information required) and also of the technical characteristics of the machine, because of the time required to exchange halos. A set-up with many tiles on many CPUs requiring large amounts of exchanges for halos may be less efficient than a setup with less tiles (each being larger). The performances quantified on a given type of machine is not transposable to another since they are sensitive to CPU types etc.*

We have added this sentence in the text: "The optimal number of CPUs for the parallelization of the transport scheme depends on the size of the tiles and also of the technical characteristics of the machine, because of the time required to exchange halos."

47) Figure 6 and 7. Is it possible to show difference plots? And more statistics (RMSEs, correlation, bias? Error reduction? Are these really surface concentrations? They are column measurements, right? What about initial conditions? Has this change as well since these are part of the control vector? Superscript on units?

in gmd-2019-186-author_response-version2.pdf p56: *Figure 7 indeed represents tropospheric columns, the legend has been corrected.*

in gmd-2019-186-author_response-version2.pdf p58: *The initial conditions are slightly changed. This is now described in Section 4.2.2: "With prior error standard deviations assigned to 15% of the initial conditions, the changes in initial conditions are very small (not shown) and do not affect the posterior emissions (test B, Figure 8)."*

I agree with the reviewer the it is difficult to tell the differences apart. Could at least a suitable colour map be used in order to visually appreciate the differences?

We thought that the added Table 4 and Table 5 with statistics for the comparison were sufficient to understand the differences between simulations and observations. We have added differences maps in Figure 5 and Figure 6.

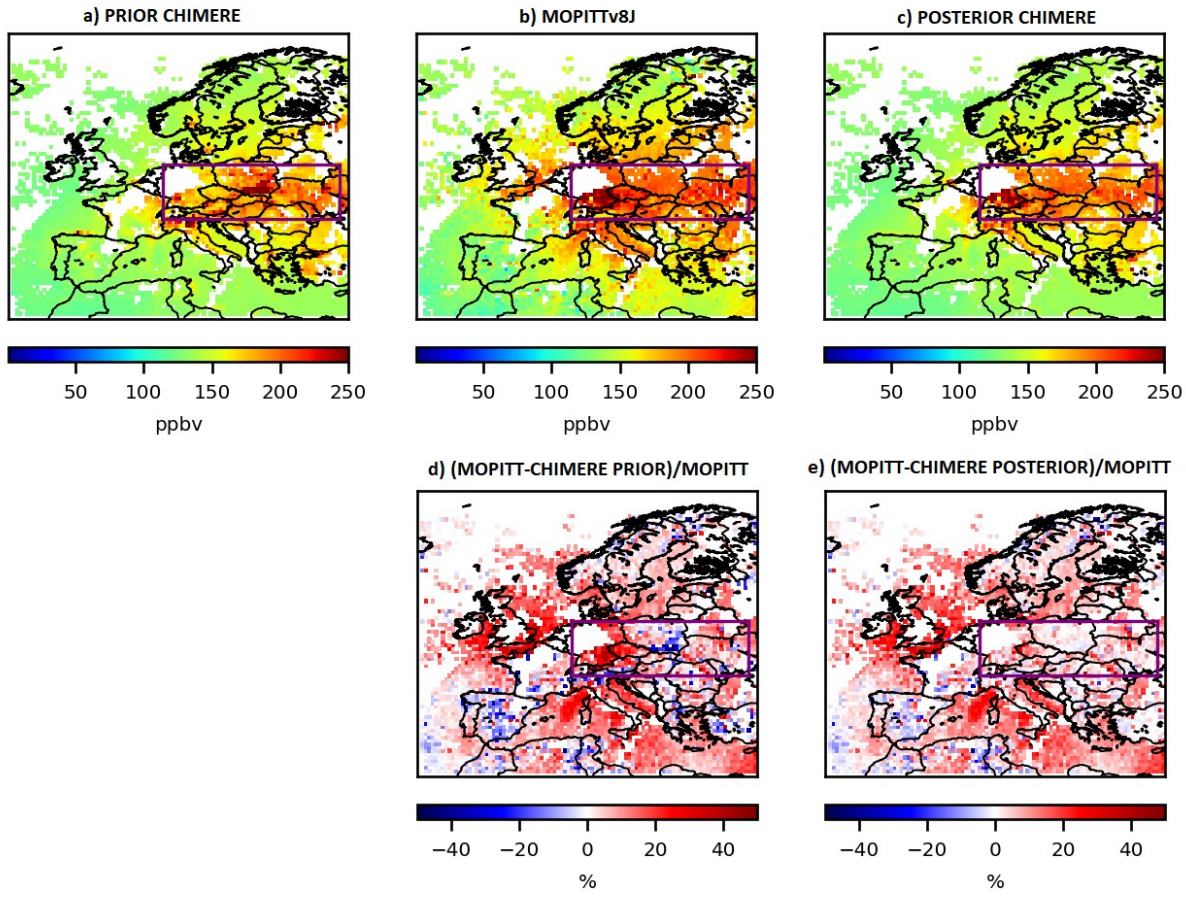

Figure 5. Mean CO collocated surface concentrations from the 1st to the 7th, March 2015 a) simulated by CHIMERE using the prior TNO-GHGco-v1 emissions and the climatological values from the LMDZ-INCA global model for initial and boundary conditions, b) observed by MOPITTv8-NIR-TIR and c) simulated by CHIMERE using the posterior emissions, in ppbv, at the 0.5°x0.5° grid-cell resolution. ***Relative differences between MOPITT and d) the prior CHIMERE simulation or e) the posterior CHIMERE simulation, in %.*** *Statistics for the comparison between simulations and observations are given in Table 4 for the area in the purple box.*

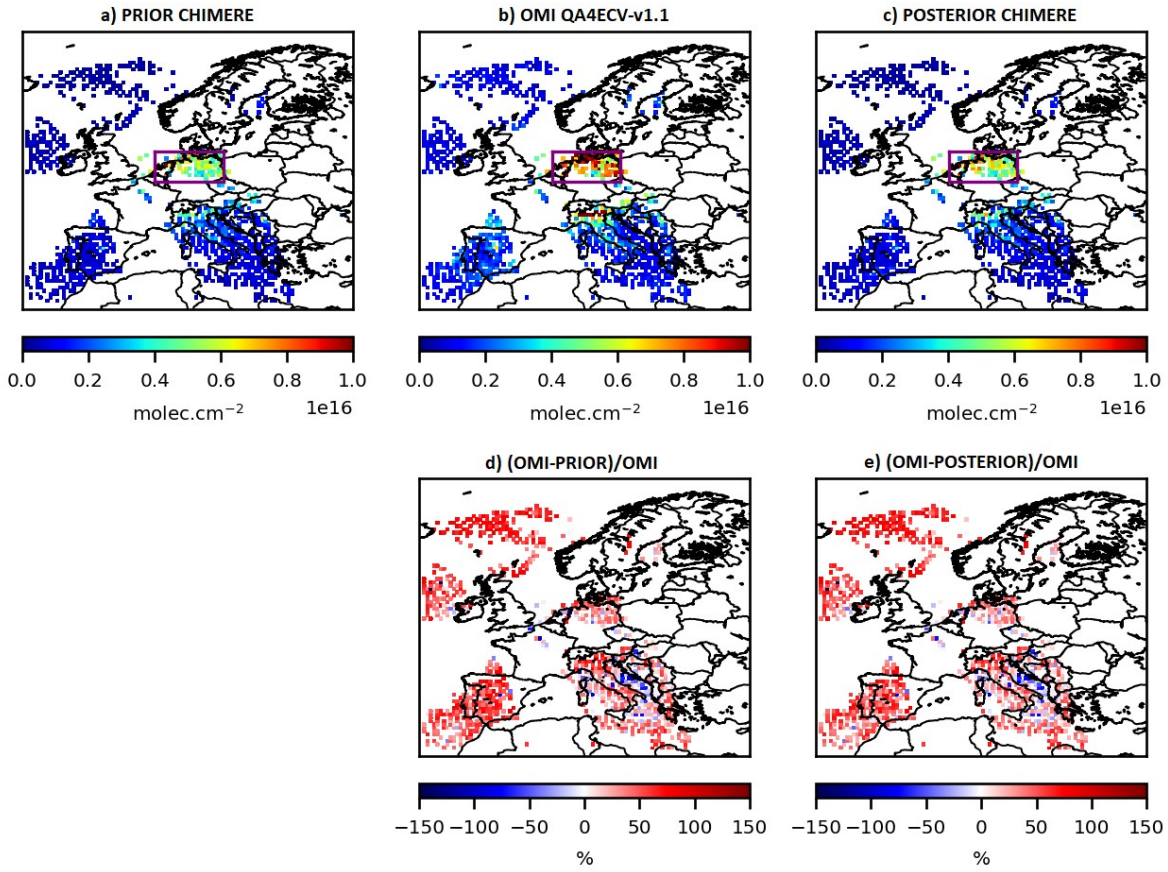

Figure 6. NO_2 collocated tropospheric columns a) simulated by CHIMERE using the prior TNO-GHGco-v1 emissions and the climatological values from the LMDZ-INCA global model for initial and boundary conditions, b) observed by OMI and c) simulated by CHIMERE using the posterior emissions, in 10^{16} molec.cm^{-2}, at the 0.5°x0.5° grid-cell resolution, the 19th, February 2015. ***Relative differences between OMI and d) the prior CHIMERE simulation or e) the posterior CHIMERE simulation, in %.*** *Statistics for the comparison between simulations and observations are given in Table 5 for the area in the purple box.*

48) Section 4.2. Should this be presented prior to section 4.1.3 since some of the plots are for the posterior estimates?

in gmd-2019-186-author_response-version2.pdf p56: *We do not agree, we have kept the sections as initially presented.*

The section 4.1.3 is about the CO Sensitivity to emissions and to initial and boundary conditions. We still prefer to keep this section before the Section 4.2 about the inversions.

49) Section 4.2.1. Can this be summarized in a table and discuss a little bit in the text as to the rational of the choice of these parameters?

The Table 1 summarizes our examples for the definition of the control vector and for the construction of the B matrix since the first review and Table 1 is still present in the last version of the manuscriptgmd-2019-186-manuscript-version7.pdf.

Section 4.2.1 and Table 1 describe the spatial and temporal resolution of our control vector. We have further improved the definition of our control vector x: "For the CO inversion, the control vector x is:

- **the CO anthropogenic emissions at a 7-day temporal resolution, a 0.5° ×0.5° (longitude, latitude) horizontal resolution, and over the first 8 vertical levels, i.e. for each of the corresponding 101×85×8 grid cells,**
- **the CO lateral and top boundary conditions at a 7-day temporal resolution, at a 0.5° ×0.5° (longitude, latitude) resolution and over the 17 vertical levels of CHIMERE, i.e. (2x101 + 2x85) x 17 grid cells,**
- **the CO 3D initial conditions for the 1st March 2015 at 0:00 UTC , at a 0.5° ×0.5° (longitude, latitude) resolution, and over the 17 vertical levels of CHIMERE.**

Considering its short lifetime, there is no boundary conditions for NO_2. For the NO_x inversion, the control vector x is:

- **the NO and NO_2 anthropogenic emissions at a 1-day temporal resolution, at a 0.5° ×0.5° (longitude, latitude) resolution and over the first 8 vertical levels, i.e. for each of the corresponding 101×85×8 grid cells,**
- **the NO and NO_2 3D initial conditions for the 19th February 2015 at 0:00 UTC, at a 0.5° ×0.5° (longitude, latitude) resolution and over the 17 vertical levels of CHIMERE."**

Am I to assume that NOx emissions are estimated only for 1 day, and all days are the same?

NO_x emissions are indeed estimated only for 1 day in our illustration, as now indicated in the beginning of Section 4: "We have chosen to present an illustration of CO inversion **over seven days,** the first week of March 2015. Considering the short lifetime of NO_x of a few hours [Valin et al., 2013; Liu et al., 2016], we have chosen to present illustration of NO_x inversion **over one day,** 19 February 2015**."**

All days could use the same inversion set-up with relevant prior emissions and observations.

For CO, what do you mean by 7-day? Average?

It is not an average, the same increments are applied for the 7-day window, as seen in the definition of the control vector.

How are emissions incorporated in CHIMERE in terms of time? Is there a distribution? i.e., diurnal and weekly cycle?

The anthropogenic emissions are constant within an hour, and the biogenic ones are linearly interpolated within an hour. The temporal distributions of the emissions are already described in the Section 4.1.2 in manuscriptgmd-2019-186-manuscript-version7.pdf: "Temporal profiles are also provided per Gridded Nomenclature For Reporting (GNFR) sector code (variations due to the month, weekday and hour)."

These points seem unchanged, has the explanation been lost? The table may be not necessary, but some questions are unanswered.

We have answered all the questions of the point 49. Table 1 is still present in the last version of the manuscriptgmd-2019-186-manuscript-version7.pdf. We have improved the definition of our control vector x, to further help the understanding of our case illustrations.

51) Section 4.2.3. Is it possible to break down the components of J?
in gmd-2019-186-author_response-version2.pdf p45: *The control vector and observation vectors gather fluxes and concentrations (respectively) over the whole data assimilation temporal window. Decomposing the cost function into a sum of comparisons between the projection of the control variables in the observation space and the observations at different times would actually make it quite confusing and uselessly complicated.* **Therefore we prefer not doing it.**

How about emission error reduction?
in gmd-2019-186-author_response-version2.pdf p46: *As already discussed in Section 2, posterior error covariances are not a straight forward product of variational inversion systems. We have added a reference to Rayner et al.[2019] in this paragraph.*

in gmd-2019-186-author_response-version2.pdf p57: *The emission error reduction is not a straightforward product of the method.***"**

How do you ensure that these increments are "resolved by the observations".
in gmd-2019-186-author_response-version2.pdf p57: *The method ensures it by design. Sensitivity tests will be done later over longer periods for example with constant emissions to quantify the impact of the observations in the system.*

It would be great to see error reduction plots, if posterior error covariances are calculated.
in gmd-2019-186-author_response-version2.pdf p58: *They aren't, as explained before.*

How about initial conditions? Did this change as well?
in gmd-2019-186-author_response-version2.pdf p58:*The initial conditions are slightly changed. This is now described in Section 4.2.2: "With prior error standard deviations assigned to 15% of the initial conditions, the changes in initial conditions are very small (not shown) and do not affect the posterior emissions (test B, Figure 8)."*

Only part of the questions are addressed in the text.
We have answered all the questions of the point 51 and we think that only 2 (about the emission error reduction and about the changes in initial conditions) should lead to changes in the text.

52) Line 508-516. What is the implication of this to overall cost and computing and optimality of minimization including error correlation of CO and NOX (and spatial correlation against superobbing) as well as increase in dimension of control vector? This also entails using this system at higher spatiotemporal resolution, right? It would be great to have a section on limitations before future implications

Part of the question is still not addressed.

We have answered above about the error correlation of CO and NO_x. We have also answered about TROPOMI high spatial resolution: "*The high-resolution imaging of TROPOMI will indeed entail using PYVAR-CHIMERE at higher spatio-temporal resolutions, but for smaller domains (i.e.,*

over countries rather than over Europe as a compromise between resolution and the computational cost.**"**

We agree to insert this sentence in the conclusion of our manuscript: "*These new space missions with high-resolution imaging have the ambition to monitor atmospheric chemical composition for the quantification of anthropogenic emissions*. **It will indeed entail using PYVAR-CHIMERE at higher spatio-temporal resolutions, but probably for smaller domains (i.e., over countries rather than over Europe) as a compromise between resolution and the computational cost**."