

Tropospheric CO vertical profiles deduced from total columns using data assimilation: methodology and validation

by :

L. El Amraoui, J.-L. Attié, P. Ricaud, W. A. Lahoz, A. Piacentini, V.-H. Peuch, J. X. Warner, R. Abida, J. Barré, and R. Zbinden

Paper under review for : Atmospheric Measurements and Techniques

Below our responses to the three reviewers:

We wish to thank the three reviewers for their excellent reviews. Their comments have been of great help and have allowed us to improve the quality of the paper, and also clarify the main objective of the paper concerning the methodology used, as well as the limitations associated with deducing the vertical profile from the total column information.

Furthermore, we take this opportunity to explain better the meaning and philosophy of this study: The total column is generally produced from the retrieved profile using a simple integration from the surface to the top of the atmosphere. The question we pose in this paper is the following: Can we use an assimilation system to deduce the CO vertical profile from the total column using the adjoint of the integration operator? If yes, under what conditions?

It is possible our discussion on the applicability conditions and the validity of this method was not clear enough in the original submission. This is made clearer in the revised manuscript, including the conclusions section.

We have demonstrated in the paper that the deduction of vertical profiles from total columns is valid for MOPITT-V3 for which the DFS (Degrees of Freedom of Signal) is relatively small when using only the adjoint of the integration operator. For other kind of data for which the DFS is greater than that of MOPITT3, the method has to be tested and validated with independent data.

Finally, and before responding to the reviewers comments, we explain below the advantages of our approach:

1 - We show that for observations for which the DFS are of the same order as for MOPITT3 (~1.5 for vertical profiles and ~1 for total columns), we can deduce the vertical profiles with good confidence compared to the retrieved profiles, and that the assimilation of total column gives almost similar results as the assimilation of vertical profiles. Consequently, for these data, it may be sufficient to assimilate the total column since its assimilation is less expensive and faster compared to the assimilation of vertical profiles.

2 - For other species for which the DFS is of the same order as in this study and for which we do not have the profiles, we can apply this method to deduce vertical profiles of these species.

Below are our responses to the comments from the reviewers.

Referee 1

General comments:

This work presents a method to derive the vertical profile of CO from its total column using data assimilation. Comparison with model free run, the authors showed that the assimilation result with MOPITT column data significantly reduced the difference relative to that with the MOPITT CO profile data. This method can be extended to the chemical species that can only be measured as the total column. For better interpretation of the assimilation result, the authors should demonstrate whether their assimilation has enough vertical sensitivity. I recommend the paper for publication after consideration of the points below.

Specific comments:

1: The zonal average degrees of freedom for signal (DFS) for the TIR-only MOPITT retrievals is typically about 1.5. When we use the column data, we only have 1.0 freedom, which will lead to the following question: do we have enough vertical sensitivity, or we are scaling the vertical profile uniformly?

➔ The method proposed in the paper does not scale the vertical profile uniformly (see our response to this referee's 2nd comment where we compare our method to the uniform scaling method).

The assimilation of CO total column in this work is done using the general formalism of data assimilation with which we can assimilate any level 2 product of the species and which consists of finding the best state of the atmosphere x^a knowing the background state x^b and the observation y . Since the observation operator H is linear, the analysis state can be expressed as:

$$x^a = x^b + K (y - H.x^b) \quad \text{where} \quad K = BH^T (HBH^T + R)^{-1}$$

The method used in this study is the variational 3D-FGAT method which consists of minimising the cost function $J(x)$ to find the analysis state x^a .

Note that in this study, although we assimilate the CO total column, the control variable is the 3-D CO field and the total column of CO acts as a constraint. In our case, y is the CO total column and x^b is the vertical profile of CO. The analysis x^a (which is a vertical profile of CO) is given by the equation above.

The update of x^a after the minimisation of the cost function is done by using : $x^a = x^b + \text{delta}_x$, where

$\text{delta}_x = BH^T (HBH^T + R)^{-1} \cdot [d]$, where d is the innovation vector. From this, we see that the correction delta_x to be added to x^a is normalised by $(HBH^T + R)^{-1}$, after it is introduced into the model space (here the CO vertical profile x) via H^T , and finally is multiplied by B .

We explain this more clearly in the revised version of the manuscript.

Figure 15 shows the assimilation can reduce the zonal mean bias well. But the bias of zonal mean could be different with regional value and could also be affected by longrange transport. As a simple test, the authors can scale the vertical profile uniformly with ratio: MOPITT_column

/ Model_column in the assimilation process. It would be very interesting to see whether the method, presented in this work, is significantly better.

➔ We have applied the method suggested by the reviewer to verify the advantages of our method. The results are given below (Figure R1). We see that the method presented in the paper is better since the results of our method compared to the assimilation of vertical profiles show smaller differences in terms of the zonal mean (Figure 15 of the revised paper) or in terms of lon-lat maps (Figure 16 of the paper) in comparison to the method proposed by the reviewer. This is particularly the case for the upper troposphere.

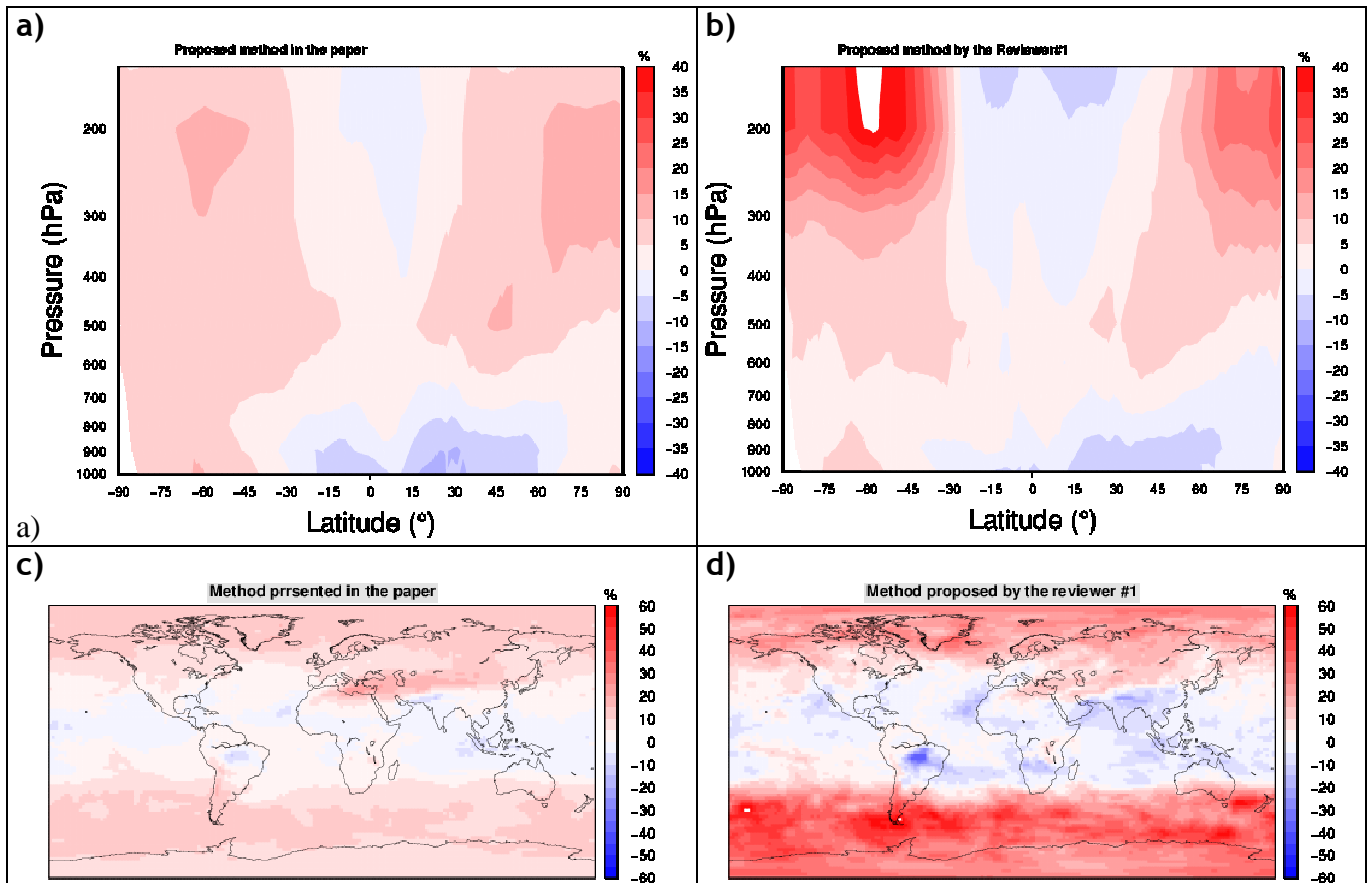


Figure R1 : (a) The difference in terms of the zonal mean between TOTCOL_ANALYSES and PROFILE_ANALYSES using the method presented in the paper. (b) is the same as figure-a but using the method proposed by the reviewer for TOTCOL_ANALYSES. (c) The difference in % between TOTCOL_ANALYSES and PROFILE_ANALYSES at the pressure level of 200 hPa. (d) the same as Figure-c but using the method proposed by the reviewer for TOTCOL_ANALYSES.

2: Page 6530, line 20-23: “The mean differences between both datasets are higher at 700 hPa than at 250 hPa. This could be explained by the reduced sensitivity of MOPITT measurements at lower levels.”

I cannot agree with this point. MOPITT V3 has weak sensitivity at surface but strong sensitivity in middle troposphere, from 700 to 350 hPa. I suspect that the difference between 700 and 250 hPa is due to the lost of vertical information by only using column data. This is why the authors need to check whether their assimilation has enough vertical sensitivity.

➔ First, we note that the redistribution of the increments after the total column assimilation is controlled by the assimilation system via the H^T operator and the B matrix (see our response to comment#1 from this reviewer).

Second, we agree with the reviewer's comment and change the sentence as follows: "this difference could be explained by the way the assimilation system redistributes the information using the B matrix. Nevertheless, this method provides better vertical profiles than the free model when compared to the MOPITT 3 profiles observations."

3: Page 6531, line 16-19: "These results are consistent with the results found by Emmons et al. (2009) comparing MOPITT CO profiles with aircraft measurements at 700 and 250 hPa. They found that the bias is larger at lower levels, being on average 25 % and 9 % at 700 and 250 hPa, respectively."

This may not be an appropriate reference. Emmons et al. (2009) found the MOPITT CO profile is biased higher at 700 hPa. However, in this work, the MOPITT CO profile is used as true value and the objective is to reproduce the "true" CO profiles by using column data. The bias of MOPITT data itself has no relation with the bias showed in Figure 7.

➔ Our point of view is as follows: In the model, there is information from emissions, chemistry and transport, and during its integration within the assimilation process, the model could bring vertical information closer to the truth even using total column assimilation. To avoid any confusion, we correct this in the new version of the paper by this sentence: 'This could be explained by the way the assimilation system redistributes the increment after the minimisation of the cost function. The information in terms of CO content given to the system is important in the lower levels of the atmosphere compared to the higher levels.'

4: In Figure 10, the deduced CO profile match very well with that from MOPITT retrievals. However, I am wondering whether it is because the regions, showed in Figure9, are too big and thus the mean profile is strongly affected by the background CO. It would be helpful to define smaller regions, for example, over Southeast Asia monsoon or CO outflow regions to observe the assimilation effect.

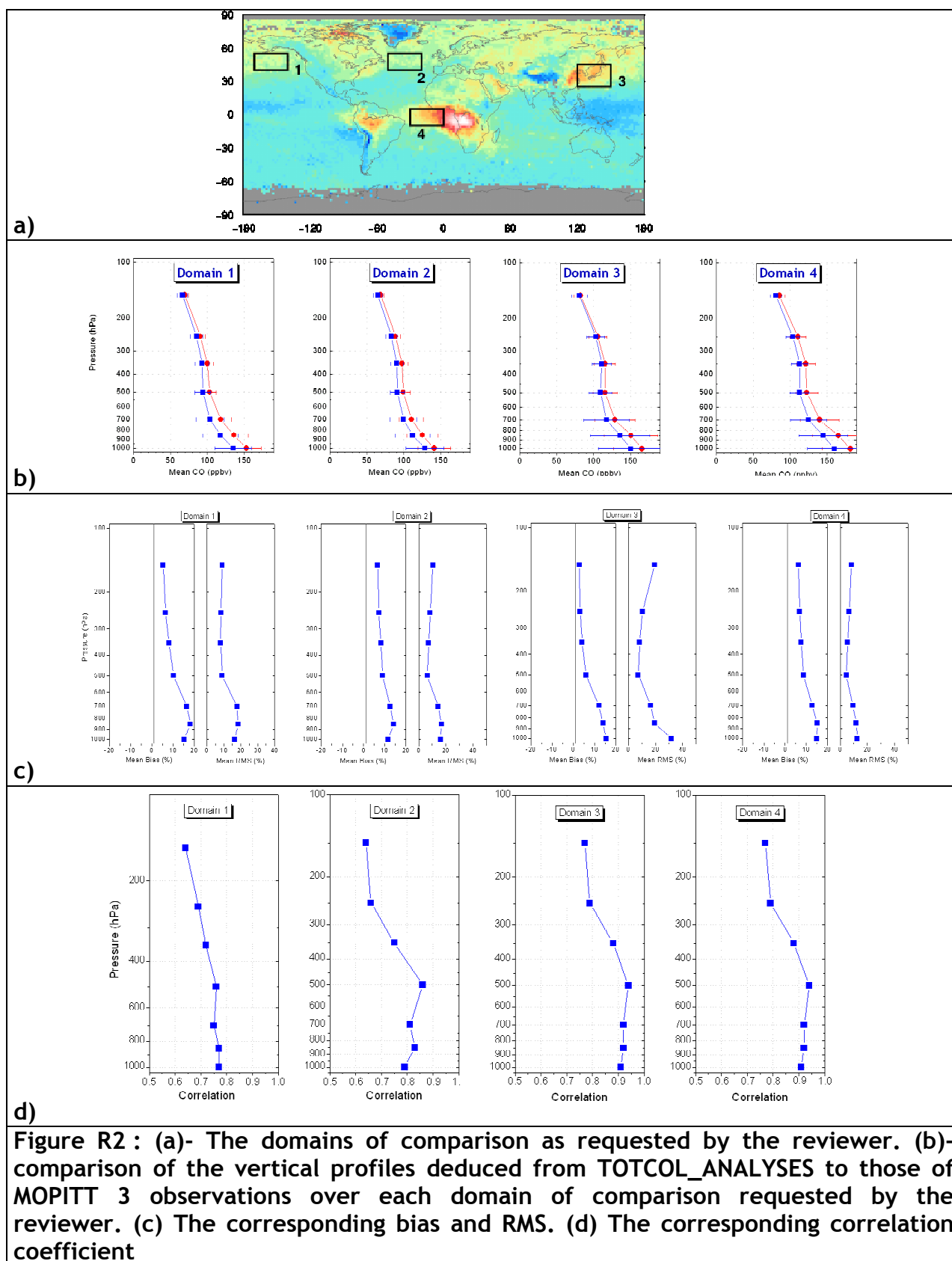
➔ Note that:

1- The focus of this paper is to deduce the CO vertical profiles at global and continental scales.

2- The resolution of the CTM MOCAGE used in this study is $2^\circ \times 2^\circ$.

These two points have to be taken into consideration for the interpretation of the results. Nevertheless, as suggested by the reviewer, we have selected smaller domains than those in the paper (see Figure R2-a below), and make the same statistics as in the paper of the vertical profiles deduced from TOTCOL_ANALYSES with respect to MOPITT3 observations in terms of vertical profiles (average, Std. Dev., Bias, RMS and Correlation). The results are presented below in Figure R2-b,c,d. Both vertical profiles, from TOTCOL_ANALYSES and from MOPITT3, are very similar and each vertical profile is found to be within the standard deviation of the other vertical profile. The mean bias for these smaller domains is of the same magnitude as for the spatial domains used in the paper. The maximum mean bias over the period of study is about 15% and the correlation coefficient ranges from 0.64 to 0.9 for all levels. These results show, again, that the vertical profiles

deduced from TOTCOL_ANALYSES are in good agreement compared to those of MOPITT3 observations even for small spatial domains.



Technical comments:

1: In Equations, vectors should be written with vector formats.

➔ Fixed

2: Page 6526, line 6: “p is the number of degrees of freedom”. Is it the best description? In 4DVAR, i should be the time step and p should represent the final time step.

➔ Variational methods exist in a variety of formulations. We use the notation of Ide et al., (1997) as suggested by Reviewer#3. We then amend the equations and improve the description.

3: The titles/descriptions of some figures are too small to read.

➔ This is due to the style of the AMTD journal which reduces the size of the figures. We have improved all figures in the revised version of the paper.

4: The description of “top” and “middle-left” in Figure 15 should be exchanged.

➔ Fixed

References :

- Ide et al., (1997), Unified Notations for Data Assimilation: Operational, Sequential and Variational, J. Meteor. Soc. Japan, 75, 181-189.

Referee 2

The main message of the paper is summarised on page 6538: "The aim of this paper is to show the benefit of using satellite CO total column data with no associated error covariance matrices and averaging kernels within an assimilation system. " After reading the paper I arrived at several fundamental objections which are detailed below. As such I cannot support publication of this paper in AMT.

➔ The main goal of the paper is presented in the first paragraph of the abstract as well as in the introduction, page 6521, line 13.

- In the abstract: "This paper presents a validation of a method to derive the vertical profile of CO from its total column using data assimilation."

- In the introduction: "The main goal of this study is to assess the benefit of the CO vertical distribution at global and regional scales."

We improve the paper and the conclusions to be more consistent with the main goal of the paper.

On page 6521 the authors write: "The proposed method has the advantage of allowing fast computation of the vertical profiles and the analyses of CO. It will be particularly useful in the future when there will be many missions providing large volumes of data for which level 2 retrievals with their corresponding characteristics (covariance matrices and averaging kernels) will be very expensive in terms of computer resources (i.e., IASI onboard METOP-A and METOP-B or future geostationary missions). " The authors argue that the removal of information stored in the kernels and covariances is valid and useful for future applications.

According to the book of Rodgers, the information stored in these kernels and covariances is non-trivial and important. It provides the details of the sensitivity profiles of the satellite observations and the uncertainty, which may vary considerably from one observation to the next related to changes in thermal contrast, albedo etc.

➔ It is never our purpose to imply that the kernels or covariance matrices are not useful. Rather, we wanted to show that the TOTCOL assimilation approach can be an alternative for these data if one wishes to provide quickly a reasonable and useful analysis, especially if the observations are retrieved operationally and are available only in terms of their total column without other associated information (e.g., kernels and covariances), as is currently done by EUMETSAT with data derived from CO IASI, where only the total column is available.

The question we try to answer in this paper is the following: for these data, can we get an idea about the vertical structure of CO from the total column? Our goal is twofold: first, deduce the vertical profiles from the total column, and second, evaluate the assimilation of CO total column compared to the assimilation of the retrieved profiles.

In the second part of the paper we compare these two analyses (total column and vertical profiles). Finally, we note again that this study remains valid for the data type for which the DFS is of the same order as for MOPITT3 and for relatively large spatial scales (global or continental).

Data volumes are huge, but this is not generally considered to be a major problem. Computer capabilities (storage, processing speed) are also growing in time. If needed, data thinning or

grouping into "superobs" may be considered. Furthermore, as shown by Migliorini (doi: 10.1175/2007MWR2236.1) there is an efficient way of storing satellite information without losing information. To my opinion such approaches are to be preferred, and the paper was not able to convince me otherwise.

➔ **We do not agree with the referee:**

- 1- First, computing resources are very different from one laboratory to another, and not all laboratories have the computing resources of, e.g., an operational centre. It should also be noted that the assimilation of the total column is much less expensive than the assimilation of the vertical profiles which takes into account their characteristics (kernels and covariances).
- 2- It is true that for local events and for shorter periods, the use of the vertical profiles with all corresponding characteristics (kernels and covariances) can certainly bring more information in comparison to the information contained in the total column. However, when looking at the global scale over long periods, the method presented in this paper shows that the TOTCOL analysis has small differences compared to the assimilation of vertical profiles, and results are very satisfactory.
- 3- For atmospheric species available only in terms of vertical profiles, this method could be used to deduce their vertical profiles if the DFS are of the same order as for MOPITT 3.

Conclusions: The authors conclude by mentioning that the approach will be applied to other satellite measurements (in particular IASI). This more or less implies a recommendation that the approach of neglecting kernels and covariances is more generally applicable. To my opinion such a statement should not be made without justification, and the approach of simplifying observations should be extensively tested for each dataset and assimilation system separately.

➔ **Again we must disagree with the reviewer.** There is no recommendation from us to ignore the kernels and the covariances. This method is an alternative way for a particular type of data and for very large spatial scales. We recall that at the beginning of this response we clarified the scope of the applicability of this method (namely, DFS of the same order as for MOPITT 3). This was clarified in the revised version of the paper. For IASI, it should be noted that this is an instrument that measures CO in the TIR (as for MOPITT), and that the DFS for IASI are of the same order as for MOPITT3. Therefore, we think that the method proposed in this paper could be applied to IASI data, especially given that CO from IASI is produced operationally by EUMETSAT using an inversion scheme based on the Neural Network method without kernels nor covariances. In future studies we will test the proposed method for IASI data.

Abstract: "Second, for chemical species that can be measured only as the total column, this method provides an attractive alternative for estimating their vertical profiles in the troposphere."

Conclusion: "The total column analyses permit us to give relevant information on the vertical structure of the CO field at global and regional scales." The authors claim that total column measurements can constrain profile shapes. This does not make sense. With one piece of information one can only constrain one aspect of a model. A local column measurement can not constrain the local profile, apart from an overall scaling.

→ Our point is that we have assimilated a typical total column observations of a measurement for which the DFS is relatively low. Note that the assimilation system used in this study is capable of assimilating any general data type. It should also be noted that whatever the type of observations assimilated by our system, the control variable of the model is still the 3D concentration. The assimilated variable only acts as a constraint on the system. After minimization, and once the minimum of the cost function is reached, the assimilation system corrects the vertical profile of the model through the adjoint of H (H^T if H linear) and by B .

In the revised version of the paper, we explain further the methodology used and we also explain that the total column acts as a constraint and that the assimilation system produces the optimal 3D increment (section 3.1).

What is a typical Degrees of Freedom for Signal (DFS) for the MOPITT CO observations used? This is crucial information but this information is not provided in the paper. When the DFS is close to $= 1$ the profile observations cannot be viewed as independent of the total column observations! Very similar assimilation results may be expected when furthermore the a-priori is not too different from the model simulated profile.

→ We fully agree with this comment. The DFS could be the key to this study, and we forgot to mention this in first version the paper. We mention this in the revised version of the paper.

Concerning the a priori, we are not sure we understand what a priori the reviewer is referring to:

- a- If it is the a priori of MOPITT3, we note that for MOPITT3, the a priori profile is constant (one single profile for all measurements) and is completely different from the model. It is also one reason for using total column from MOPITT3..
- b- If it is the model a priori, we note that the model gives different results from those of the assimilation results (see results of section 5 and the related figures).

Concerning the chi-square analysis: I do not see how this can uniquely constrain the observation error. This is only possible when B is accurately known (or known to be small compared to R). There is no proof of this in the paper. Normally, the chi-square test is used to constrain B based on a fixed R provided in the retrieval dataset.

The background covariance B optimisation and structure should be discussed in much more detail for the reader to appreciate the meaning of the results. The B matrix will determine how the total column measurement affects the profile shape. Because chi-square is a single number it can at best constrain only the average error, and does not provide information on the variability from one measurement to the next.

→ In the application considered in the paper, the observational errors are smaller than the background errors ($||R|| < ||B||$).

We do not fully agree with the comment that B must be small (strictly speaking, the matrix norm of B) compared to R .

If the observation errors are smaller than the forecast errors, the information brought by the observations with respect to what was already known is considered as a “signal” and the analysis could be better. In the reverse case ($||B|| < ||R||$), the information brought by the observations with respect to what was already known is considered as a “noise” and the analysis could be poor. For more details, see: ([Zupanski et al., 2007](#)).

- The Chi2 test in data assimilation gives an indication of the consistency of the covariance matrices B and R ([Lahoz et al., 2007](#); [Ménard and Chang, 2000](#)). A Chi2 value close to 1 is a good indication of the consistency of the assimilation algorithm ([Talagrand et al., 2003](#)). Therefore, if B is known, the Chi2 test can give information about R. We explain how the B matrix is parameterized in the revised version of the paper.

How large is the variability of the observation errors in the retrieval product? Is this variability reflected in OMF? This aspect is not discussed.

➔ We are not sure what the reviewer means by retrieval product? Does the reviewer refer to the total column or to the vertical profiles?

- If this concerns the errors of the total column, we recall that these data are considered in the assimilation system without the error retrieval. In a first step we estimate these errors and afterward we assimilate these data.
- If this concerns the errors of the vertical profiles, we recall that these errors were taken into consideration during the assimilation of the vertical profiles. However, this assimilation product is considered as the true state during the validation of TOTCOL analyses and we have not evaluated the analyses from the vertical profiles.

For other data assimilation approaches with an other chemistry model and a different formulation of the B matrix there may possibly be much larger differences between the column and profile analyses. I am not at all convinced that the results can be generalised to other assimilation systems and other satellite datasets.

I would like to stress that the objections given above concern the conclusions and objectives of the paper. The MOCAGE-PALM assimilation system seems to be working very well, producing high quality analyses.

➔ We agree with the reviewer that the method cannot be applied to all data. This is a detail that we forgot to mention in the paper. We clarify this in the revised version of the paper.

In addition I have a few related detailed remarks:

To my taste there are too many plots in the paper. All the details on the global distributions, zonal means, regional means and vertical profiles are not needed for the main conclusions, and the results sections may be shortened.

6522, 125: What is a typical value and range of DFS ? This is crucial information.

➔ The typical DFS of CO MOPITT3 is ~1.5 for vertical profiles and ~1 for total columns. This information is now added in the paper.

6527, 120: "Chi-square close to 1 indicates consistency between both error covariances." I would dispute this. It shows that the sum of both covariances is consistent with the observed OMF.

➔ Conventionally, the normalized OMF gives a notion of the validity of assumption of a Gaussian distribution of the errors, and the Chi2 test is a self-consistent diagnostic which gives an indication about the consistency of the assimilation algorithm, see for example:

- Talagrand (2003), (page 93) : "...This test provides a very simple diagnostic of global consistency of an assimilation algorithm. »

- Lahoz et al., (2007), (page 5761) : "...This discrepancy can arise from an incorrect estimate of B or R. »

- Ménard and Chang, (2000), (In the abstract) : « A robust chi2 criterion, which provides a statistical validation of the forecast and observational error covariances.... »

6528, fig 1: First I would like to see evidence that B is well specified. The optimal value of R will depend on the construction of B.

➔ The description of B is now given in the paper. Its specification within the assimilation system MOCAGE concerning MOPITT3 CO is based on many studies: (e.g., El Amraoui et al., 2010) where the MOPITT3 analysis in terms of vertical profiles gives good results compared to independent data such as AIRS and MOZAIC.

6533, fig 10: It is important to know what the (average) DOF is in these cases. If this is small there could be a lot of a-priori information mixed in, and a good comparison of the profiles is not very meaningful.

➔ This information is now added in the paper.

6535: It would be interesting to know why the assimilation of total columns and profiles (with kernels, covariances) gives very similar results. Are the a-priori profiles used in MOPITT similar to the MOCAGE profiles?

➔ MOPITT 3 has a single a priori profile over all the globe, however the CTM MOCAGE has profiles that are different from this a priori profile.

Both analyses are very similar because the information brought to the assimilation system by the vertical profiles and by the total column are almost the same.

References :

- El Amraoui, et al., (2010), Midlatitude stratosphere-troposphere exchange as diagnosed by MLS O3 and MOPITT CO assimilated fields, Atmos. Chem. Phys., 10, 2175-2194,

- Lahoz et al., (2007), Data assimilation of stratospheric constituents: a review, *Atmos. Chem. Phys.*, 7, 5745-5773
- Ménard and Chang, (2000), Assimilation of Stratospheric Chemical Tracer Observations Using a Kalman Filter. Part II: x2-Validated Results and Analysis of Variance and Correlation Dynamics. *Mon. Wea. Rev.*, 128, 2672-2686.
- Talagrand (2003), A posteriori Validation of Assimilation Algorithms (pp. 85-95), in: R. Swinbank, V. Shutyaev and W. A. Lahoz (editors), *Data Assimilation for the Earth System*. Dordrecht, The Netherlands: Kluwer Academic Publishers, Nato Science Series.
- Zupanski et al., 2007: Applications of information theory in ensemble data assimilation, *Quart. J. Roy. Met. Soc.*, 133, 1533-1545

Referee 3

The main subject of this paper is a comparison of total columns and vertical profiles as observational input for data assimilation of CO. The paper is a struggle to read because it provides excessive detail on uninteresting things while giving little or no attention to the real issues. For example, the specification of the observation error is not controversial or new, so it could have been covered in a few sentences. At the moment it requires an entire section, one table and three complex figures. In contrast, the specification of the background error covariance matrix is crucial to the data assimilation problem, particularly for assimilation of total columns, yet it is not described.

The main result of the paper is that assimilation of either columns or profiles gives similar results. That contradicts an awful lot of research over the years, as well as common sense, so the unanswered question is: why does this happen? The answer may lie in the way the vertical correlations are specified in the background error matrix (B), or in the information content and number of vertical degrees of freedom available from the retrievals - this comes from the averaging kernel matrix A. Neither of these aspects are properly covered in the paper. To understand what is going on, the authors need to look at how the A and B matrices combine to project information from the observations (column or profile) into model space. Ideally they need to compute this explicitly, but if this is difficult they could show examples of the vertical patterns of increments resulting from the assimilation of single observations on single levels.

➔ We conducted the test requested by the reviewer by assimilating a single observation for each kind of dataset: first, a single total column with its corresponding error; second, a single vertical profile with its kernel and covariance. The single observation for both datasets corresponds to an observation from 30 August 2008 at 02:19 (lat=39.08 and lon=-60.69).

Note that for this test, we start both assimilation experiments from the same initial field corresponding to that of 30 August 2008 @ 00:00. This later field has been obtained using the assimilation of CO MOPITT3 vertical profiles from the beginning of August (1st August 2008). This means that this initial field is considered as a good state of the atmosphere at the beginning of this assimilation test. The information content to be added by a single observation to this initial field should be very small, especially considering that this single observation is made during the first time window of the assimilation.

We then conducted two separate assimilation tests for both the total column and the vertical profile starting from the same initial field.

Note again that the control variable of the model is the 3D increment. This means that for both tests, the increment is spread over each grid point of the model.

Below, we present the behaviour and the shape of the respective increments at different levels in terms of lon-lat maps as well as their vertical structure.

For each level of the model, we see that the increments are Gaussian around the point of the measurement for both analyses. This is consistent with the use of the matrix B.

Concerning the assimilation of the total column, the vertical profile of the increment shows that its maximum is located in the lower layers of the troposphere.

Concerning the assimilation of profiles, the vertical structure of the increment shows that the maximum effect is located in the mid-troposphere.

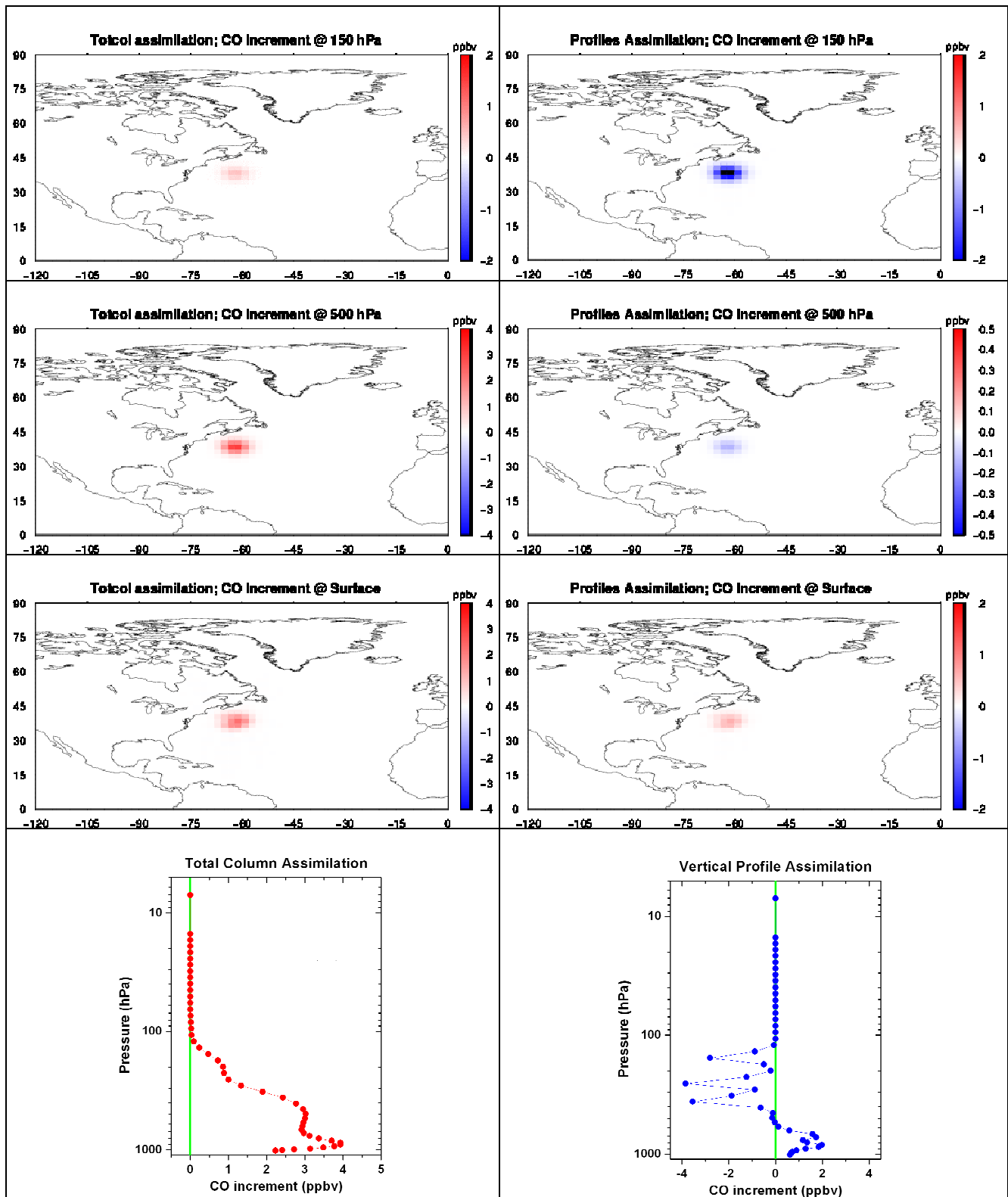


Figure R3: Behaviour of the increment for one single observation assimilation: Total column (left) and vertical profile (right). 1st line: both increments at 150hPa; 2nd line: both increments at 500hPa, 3rd line: both increments at the surface. Bottom line: the vertical structure of both increments at the observation location and over the model vertical levels.

In summary, I would like to see a lot of the current paper compressed, to be replaced by a major effort to properly explain why assimilation of columns or profiles makes little difference to the results.

Major points

1) p6521, lines 15-26, provide the justification for this work, which seems weak. It seems to imply that assimilation of columns is a necessity to avoid computational performance problems. However, many groups assimilate profile data from multiple satellites in an operational context, for example the MACC project. The real justification, perfectly valid, is revealed only in the conclusion: the idea is to assimilate CO columns from IASI, where presumably there is not enough information to provide a vertical profile.

➔ **The assimilation of total column is not needed to avoid the problems of storage and computing, but it is clear that the assimilation of total column is much less expensive than the assimilation of vertical profiles which takes into account all corresponding characteristics such as the kernels and the covariances. Moreover, it should be noted that many research centres and/or groups do not have the computing and the storage resources of, e.g., ECMWF.**

We explain in the revised version of the paper that this method is valid for observations for which the DFS is of the same order as for MOPITT3 over very large scales (global or continental). Consequently, the method presented could be an alternative for producing global CO analyses fields. But it is clear that for local problems, the use of profiles with all other characteristics (kernels, covariances) is a more natural way for producing the analyses and document the local events.

2) The description of the DA system in sections 3.1 and 3.2 lacks some essential details, principally a description of the B matrix, and it seems confused in some of the notation (that will be covered in minor points).

➔ **We correct the text and now provide a proper description of the equation.**

3) In section 2, the chi squared method is presented in excessive detail yet it ignores the real issue. The time variation of chi squared would be uninteresting for one of the three data categories (land day, land night, ocean) let alone all 3. It would be enough just to present table 1 without showing figures 1-3 at all. The real issue is that the chi squared method assumes the background error variances are correct in magnitude. The quantification of the observation errors makes sense only if this is true.

➔ **The matrix B is explained in the text. Regarding the figures, we think that it is important to see how the chi2 test behaves during the assimilation process with respect to the time. This verification is crucial because it is the only criterion we have adopted to specify the errors of observations in terms of total column. For this reason and following the recommendation of the reviewer, we kept only one figure among the three.**

4) Sections 4 and 5 could easily be compressed to make for a lighter and more interesting read, and to make room for an explanation of what causes the results. For example, figures 5, 6, and 7

could be combined. There is no need to always split the observations into the different groups of land/day, land./night and ocean. The authors should only split things out when it reveals a feature of real interest.

➔ **We have made a great effort to condense the paper, especially for these two sections (4 and 5).**

Minor points

1) Abstract. This is over-detailed (is there any need to give details of the chi squared errors?) and does not need to include labels like "LAND_DAY", "LAND_NIGHT" and "TOTCOL_ANALYSES" and "PROFILE_ANALYSES". These capitalised labels are in general a bit of a distraction and make the paper harder to read, not easier.

➔ **We have made an effort to reduce the length of the abstract.**

2) Page 6519, line 23 to 6520 line 5 gives a long list of instruments and acronyms that have nothing much to do with the current paper. Such a long list is quite hard to read. It would be more accessible if it were given in a table, but really I don't think it is necessary at all.

➔ **We have removed this list of instruments.**

3) Page 6520, lines 6-14. A list of the constituents retrieved by IASI and a description of the METOP satellites is pointless in the context of the current paper.

➔ **We have removed this list of constituents.**

4) Page 6523, lines 10-14. This very long list of papers is not useful for understanding the current work; it should be pruned or removed.

➔ **We have removed this list of papers.**

5) Page 6524, lines 13-15. Again, this is a long list of citations that brings nothing to the current work. Things should only be cited if they have a direct relevance.

➔ **We have removed this list of citations.**

6) Page 6525, lines 8-18. Again, a list of citations that adds nothing to the understanding of the current work, and should be removed.

➔ **We have removed most of these citations.**

7) Section 3.1: notational problems and confusing description:

a) In equation 1, do you really mean $H(x_i)$, not $H(x)$? In other words, where does the forecast model come in, how is information propagated from the analysis time to the observation time?

➔ **We give the description of the equation following the notation of Ide et al. (1997) as suggested by the reviewer (see below).**

b) Is p really the number of degrees of freedom? Conventionally $i=1,p$ represents the different time-steps in the model. See Ide et al. (1997, J. Meteor. Soc. Japan) for standard DA notation conventions.

➔ The notation of all the equations is now fixed in the revised version of the paper (see the comment (a) before).

c) Equation 2 seems to have lost some subscripts compared to equation 1. Where have they gone? The text should explain.

➔ We correct the equation.

d) p. 6525 line 21 to 26. It's not clear what these lines are attempting to say. Whatever it is, it could perhaps best be explained with mathematical notation, or else with a much more carefully explained piece of text. As I read it now it is wrong, anyway. The increment is not just a projection through H^T but also through the background error covariances.

➔ In the revised version of the paper, we explain better the methodology and how the total column assimilation works within our assimilation system and how the TOTCOL_ANALYSES impact the CO vertical profile of the model.

8) Section 3.2, as mentioned before, a section on "error specification" needs to include the background errors.

➔ We add a section describing the background error covariance matrix (see section 3.3)

9) p. 6529, line 16; "this indicates .. the bias .. is reduced". Standard deviation says nothing about bias, this cannot be true as written.

➔ We remove this last sentence.

10) Section 4.1 seems to have a strange viewpoint in which the analyses provide the reference against which the MOPITT vertical profiles are being validated! How can the authors justify this, given that the analyses are a combination of CTM and total CO observations? Are they saying that the CTM is more accurate than observations?

Examples:

a) Page 6530, line 21: [biases in the lower troposphere] "could be explained by the reduced sensitivity of MOPITT .. at lower levels".

➔ This is the same remark as made by reviewer#1; this has been fixed as: This difference could be explained by the way the assimilation system redistributes the increment. The information given to the system from this increment is important in the lower levels of the atmosphere compared to the high levels.

b) Page 6531, lines 4-7 similarly

➔ Fixed

c) Page 6531, lines 16-19 similarly

➔ Fixed

Really, the biases between the analyses and the MOPITT vertical profiles are a diagnostic of how the assimilation system is working, not an absolute verification of the quality of MOPITT. As an aside, some independent validation of the CO fields would have been nice in this paper; everything is self-circular at the moment.

➔ We add a section (see section 5.4 of the revised version of the paper) where we validate TOTCOL_ANALYSES and PROFILE_ANALYSES with in-situ MOZAIC independent observations.

11) Page 6532, lines 21-23. "the spatial extent..." This sentence doesn't make sense.

➔ We replace this sentence by: the spatial distribution.

12) Page 6532, lines 17-28. This introduction/justification seems a bit excessive; it could easily be cut down.

➔ We condense the introduction of section 4.3.

13) page 6533, line 12: Vertical profiles "agree within the standard deviation of both datasets". This is spurious; the t-test for the difference of two populations is based on the standard deviation divided by $\sqrt{\text{population size}}$, a much smaller number. Statistically, these populations are probably very different; this does show in the rather low correlations exhibited in Figure 12.

➔ Our aim is to show that both datasets are very similar and each profile is situated within the variability of the standard deviation of the other profile.

14) page 6533 line 17: "mean RMS" - this is confusing and needs some explanation

➔ We mean the corresponding RMS. We change it in that way.

15) page 6534 lines 5- 15: analyses and observations are "generally in good agreement"
- this is unconvincing, given the low correlations in Fig. 12. Also, this section is a further example of the curious attitude where the analyses are seen as verifying the MOPITT retrievals. Here, the lack of sensitivity of MOPITT to the lower troposphere is blamed. This is another place where a look at the A and B matrices would help illustrate what the authors are trying to say.

➔ We do not understand this comment from the reviewer.

References :

- Ide et al., (1997), Unified Notations for Data Assimilation: Operational, Sequential and Variational, J. Meteor. Soc. Japan, 75, 181-189.