

Interactive comment on "Global distribution of CO₂ in the Upper-Troposphere and Stratosphere" by Diallo et al.

Diallo, Legras, Ray, Engel and An  l

Correspondence to: m.diallo@fz.juelich.de

Dear Editor,

We have made substantial changes to the manuscript in order to seriously consider the reviewers suggestions and comments. These major changes include: Three new figures that highlight the uncertainties on the comparison between SOLVE and model reconstructed CO₂, the CO₂ distribution from the upper-troposphere into the stratosphere as well as the uncertainties that bound the CO₂ product.

Best regards,
Mohamadou Diallo (on behalf of the co-authors)

Answer to anonymous referee #1

We thank referee #1 for his comments and suggestions. Comments by the referee are highlighted and followed by our answers.

Although the paper contains some interesting material which should be published, the manuscript itself has poor quality in many parts. The preparation of the manuscript was not done with the necessary care. Several paragraphs and sections need substantial revisions! The structure of the paper is fair but the grammatically preparation is superficial and the discussion with regard to the scientific content is inadequate. Particularly Chapters 5 and 6 are unclear and both do not contain a detailed discussion of the data product. In the following here are my major points and general concerns:

Major points:

- 1. In Section 3.1, I am missing a critical discussion of the boundary conditions for your exercise, for instance regarding the used data (i.e. the 1989-1999 time period from ground stations): what are the uncertainties of the calculated trajectories (e.g. with respect to the linear interpolation along latitudes (page 8, line 15-16))? Why are the CO2 values from the CarbonTracker considered only at 5 km altitude? Are the results of TRACZILLA affected by the assumed initial conditions? How reliable are the TM5 results in comparison to other CTMs? At least it would be necessary citing relevant literature to motivate your approach this section does not show any reference.*

It's not straight forward to estimate of the initial condition in the 1989-1999 as there's no reference global distribution of the CO₂ concentration. The linear interpolation is meant to avoid artificial discontinuities in the assignment. The basic uncertainties in the trajectories arise from errors in the advection fields that can be sensitive in some limited circumstances, as illustrated in the comparison to balloons and aircraft measurements. It is proven (see Legras et al., 2005 and Diallo et al., 2011) that the reconstructions are weakly sensitive to the duration of the integration and to the small-scale patterns of the data used in the initialization. Only large-scale bias would be seen in the reconstructed

CO2. The CarbonTracker data are used as a boundary conditions at 5km as it is a convenient way to initialize the trajectories as they cross this surface which lays entirely in the free troposphere. The representation of vertical transport by diabatic heating and the mere usage of potential temperature as a vertical coordinate is highly unpractical closer to the surface. This is better handled by a tropospheric transport model like TM5, now properly referenced, but the details of which are highly irrelevant for our purpose. CarbonTracker data is mainly used as a low pass filter which damps the daily and local fluctuations of CO2 at the surface and carries the low frequency and large scale ground signal to the 500 hPa surface in a matter of days. A precision, we don't use TM5 data but the data from the assimilation system CarbonTracker which is coupled with TM5 for the transport.

2. *Regarding Chapter 4, I am missing at least a final statement (paragraph) discussing the uncertainties with respect to the final data product. For example regarding the point mentioned on page 13, lines 18-19: After the backward integration of few months, then the air particles of the diffusive run are fixed by using the ages calculated from the global initialization. What does this mean? How reliable are these assumptions? What is the range of uncertainties here? Again, no reference is cited in this connection. Similar on page 14, line 7 you stated: . . . with a diffusivity parameter k equal to 0.1 It would be helpful to provide more information and a critical discussion of the assumptions used here.*

This paragraph that contained an error has been rewritten. The method used here is based on previous works (Legras et al., 2005; Pisso and Legras, 2008) based on long-lived tracers which provide constraints on the diffusivity. We refer the reviewer to these previous works which contains a thorough discussion of the sensitivity of reconstructions to the diffusivity.

3. *Chapter 5 is really poor. The description and explanation of results in Section 5.1 is very vague and need a substantial revision. This kind of evaluation (i.e. a comparison of colored lines) presented here is not sufficient. For example (page 14, line 24): The large peaks seen, e.g. . . . are due to dives and the discontinuities on . . . are also due to*

fast altitude changes. It is difficult to understand what you are meaning. In particular the mentioned flights are not fitting together, some of them are missed . . . and by the way, Figure 2 is not mentioned at all in Section 5.1. Figure 2 is showing some discrepancies? What are the specific reasons? What is the grey shaded area standing for in Fig. 2? What is needed here is a detailed description of the data and a critical discussion of the strength and weaknesses. Your final statement (page 15, line 11) ... explains very well the measurements ... is not sufficient.

This section has been heavily rewritten

4. *I have the same concerns with respect to the other sections in Chapter 5. In the beginning of Section 5.2 a detailed comparison is announced, but it is definitely not given in the following. It is still a description of pictures. And it is a mess: Four balloon flights are presented in Fig. 3, three of them are showing a good agreement (but only two of them are mentioned in the text; page 15, line 19); one shows significant differences!? The explanation at the end (lines 22-23) is ... the errors of the reanalysis which are not accounted in our statistical test might be large enough to explain the shift. This is definitely not an adequate explanation!*

We have improved the discussion. The third case is associated with a meteorological front, a situation which is well known to amplify the advection errors as previously shown by, e.g., Pisso and Legras (2008).

5. *In Section 5.3 you said that you choose the latitude ranges 10S-20N and 50-60N. What is the reason for chosen these bins? Why are the latitudinal bins different? A more detailed discussion is necessary. Again a statement like The origin of this discrepancy is unclear, but is perhaps due to the fact that previous studies . . . By the way: which studies (no references are given)? And at the end of this section the sentence (page 17, lines 5-7) The large discrepancies ..., as it will be confirmed shortly. is not comprehensible to me.*

The reason of the choice of the two latitude bands is to select two regions respectively representative of the tropics and of the extra-tropics. Then the altitude ranges are dictated

by the location of the tropopause in these two regions, respectively at 10-11 km and at 17-18 km. We have improved the discussion and added references.

6. *What I have said in general with regard to Chapter 5 is also valid for Chapter 6. A more detailed discussion of the results and their evaluation are required. Some particular points with respect to Chapter 6: You are discussing results of the year 2010 only in Sections 6.1 and 6.2. What are the other 9 years showing (regarding annual cycle, variability, fluctuations, etc.)? Are they identical (similar) or different? Is 2010 exemplary for the other years? Beyond that, no results are showed in Section 6.2 (not either in other sections) regarding the upper stratosphere (all figures show results at altitudes below 35 km). With regard to the Conclusions you stated (page 21, line 3): In the deep stratosphere, we have found This remark is not supported by investigations presented before.*

The results of the year 2010 is considered here as exemplary for the other years. The annual cycle of the zonal mean CO₂ for the other 9 years look similar but differ to a growth rate that changes year to year. Due to this changing growth rate, the variability of the stratospheric circulation is better studied using the age of air as in Diallo et al. (2011). The 10-year average CO₂ profiles after removing the mean from the CO₂ are described in section 5.3. We have extended the top level to 42km in Fig. 6b.

7. *A separate discussion Chapter at the end is necessary, especially an assessment and classification of the data set regarding the current knowledge. In particular to point out the new findings and a rating about the quality of the provided data set would be helpful. As announced in the Introduction (page 6, line 21) The global (!) distributions of CO₂ and its transport in the UTS (!) are investigated . . . Therefore, it would be nice to see some global charts, covering the entire stratosphere in this paper: The final sentence of your manuscript (page 21, lines 12-13) makes me curious, but no respective illustrations are shown in this paper. Could you please show (and discuss!) some examples, not only to motivate using this comprehensive data set?*

The final section has been rewritten to emphasize the results. We do not fully understand

the request of the reviewer as Fig.6 is an example of the global charts which are provided from our latitude x altitude product.

Minor points:

Many acronyms are not explained, in particular in the Introduction.

These acronyms are common. Avoiding them will remove the message that we want underline here and it keep the sentence shorter as well.

Page 3, line 19: . . . increase global change . . .

Done

Page 4, line 14: . . . observations that have high resolution and are very localized, . . .

Done

Page 5, line 29 and in particular on Page 7, line 3: TRACZILLA is not explained (at least references are required).

Done

Page 10, lines 14-17: need a Revision

Done

Page 10, line 22: It is assumed that the vertical mixing is fast . . . Why? Reference is needed? How fast? Please explain.

It's assumed the vertical transport is fast mostly driven by convection.

Page 14, line 7: Units should be corrected.

Done

Page 15, lines 15-16: . . . of four middle latitude stratospheric balloon flights . . .

Done

Page 20, lines 4-5: . . . based on 22 years of data . . . over the last 11 years Sentence should be revised.

Done

Page 21, line 3: In the deep stratosphere we have found . . . no results are shown or discussed in detail.

Done

Figure 4: the chosen latitudinal bins do not exactly correspond to the coordinates of the stations providing the measurements.

Done

Figure 6: the box in the right figure (top right) says 2000-2011; should be 2000-2010.

Done

Answer to anonymous referee #2

We thank referee #2 for his comments and suggestions. Comments by the referee are highlighted and followed by our answers.

1. *The global product requires uncertainty estimates/bounds and a description of that derivation. This is essential.*

We have estimated the uncertainties of the global CO₂ distribution and added a subsection which describes the deviation (see Fig.6c ad sect. 5.3).

2. *The manuscript needs to clarify the product for global CO₂, that it is 2-dimensional (varies with latitude and altitude but not longitude). This was unclear earlier in the text, and only clarified in the final conclusion sentence. How this was generated in terms of where the receptors for the trajectories were distributed in the stratosphere is unclear. How are the initial air parcels distributed in time and space, and are they run from different longitudes and the results averaged in zonal means?*

The reconstructions is indeed produced by averaging over all parcels launched on a latitude circle. Further grouping in latitude bins is made near the pole. This is explained in section 3.1. and 3.2. The reconstructions of balloons and aircraft measurements at given locations is based on multiple launching from the same point and diffusive trajectories as explained in section 3.3.1

3. *Grammar and spelling should be checked throughout the manuscript. I noted many of the errors below, but it should be thoroughly proofread. Many sentences are awkward.*

The quality of the writing seems to deteriorate even more in the final few pages, with basic typographical errors and usage of words that are not words.

This aspect has been hopefully improved.

Specific comments:

Overall, the journal editor can comment on whether italics are or are not appropriate where used in this manuscript, generally when defining a term or acronym (for example "tropical pipe" is italicized throughout the paper).

The italics term have been removed from the text.

L21: decreases with altitude... is nearly constant with altitude - contradictory. Perhaps the authors intend that CO₂ decreases with altitude from the UT to the S but is nearly constant with altitude above 35 km?

Yes that we meant here! The CO₂ concentration decrease up to 35km and above it is nearly constant.

Page 4: L 6: where recently assessed from balloon-based

L 28, qualitative? Perhaps better worded would be "shown qualitatively good agreement with in situ observations..." Last line and L1 of Page 5, awkward sentence

We have rephrased this sentence

Page 5:

L25 grammar

L18 this diffusivity effect

L29 models

I dont completely understand why Lagrangian models would not be subject to these problems if the underlying transport model is flawed.

Grammar corrected

CTM are usually run at low resolution due to the numerical cost of chemistry. They are therefore highly diffusive and can generate spurious diffusive transport and mask barrier effects due to the dynamics. Lagrangian models are purely advective and not bounded by diffusion. Added diffusion is however necessary to represent missing turbulence in the advective winds but it is several orders of magnitude smaller than the spurious numerical diffusion of CTM.

Page 6: L10: by scarcity are the authors referring to the scarcity of CO2 observations? confusing sentence, possibly because of wording /grammar errors.

L15: is the ERA-interim analysis used in any of the previously referred-to flawed CTM models that were unable to correctly capture transport?

L15 ERA-Interim definition should be moved up here from Line 20-21.

This paragraph is rearranged and improved.

Page 7: L2: the Lagrangian transport model TRACZILLA (Legras ..), a modified version....

L5: I see now that the Lagrangian model is calculating its own vertical motions, unlike perhaps the CTMs mentioned above, and that is why it performs better?

Line 15: Please clarify the assumption being made here: this corresponds to the assumption that the CO2 in the troposphere (i.e. below this boundary condition) is constant?

The vertical motion is derived from the archived heating rates of the reanalysis. This procedure is also quite common among the stratospheric CTM. The heating rates allow to better separate the horizontal (isentropic) motion and the vertical (cross isentropic) motion. They are also smoother, being accumulated quantities, than the vertical velocities which are instantaneous sampling (undersampling actually) of fast varying fluctuations. As 500 hPa is taken as the lower boundary of the domain, trajectories starting below (as it may happen only in high surface pressure regions) are discarded.

Section 3.1, page 8. The use of clean-air data at the ground, no matter how clean the site, is a bit worrisome. Since Carbon tracker has also been used for the later years, how would it impact your results if you used a lower or higher CarbonTracker level? The 5km level is not only above the PBL, it is significantly higher than the other stations that you use in 1989-1999. It would be worthwhile to investigate the vertical gradient in CarbonTracker to see what kind of error you have just by choosing one level and assuming it is constant. Also it would be useful to look at CarbonTracker residuals against the NOAA North American aircraft network data at altitudes above the PBL, to see if it is a realistic depiction of CO2 mole fractions in the free troposphere and upper troposphere (those profiles go to 8 km).

<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/profiles.php>

If residuals are small, then using the CT gradient might be a good way to determine the uncertainty on the assumption and the choice of only one level. CT can also be used to investigate any longitudinal errors/differences.

It is not clear if the CarbonTracker mole fractions are considered as an average over all longitudes, or for the specific grid cell where the back-trajectory initiates.

L22-23: This is not a great description. CarbonTracker assimilates CO₂ observations from atmospheric stations and optimizes underlying fluxes from the listed sources (ocean fluxes, biosphere fluxes, fire and fossil fuel).

The reconstruction procedure is well posed by using a given pressure surface as a lower boundary. As far as vertical transport in the lower troposphere occurs in a matter of days, this is not going to impact our results. Taking the CarbonTracker data at 5km filters out the ground fluctuations and allows to use a fairly limited sample of trajectories to provide a smooth reconstruction.

Page 10: L19: I repeat my question from before, which could be addressed here - it is assigned based on the lat/lon of the boundary crossing after 2000 but prior to 2000 it is only the latitude that matters?

Yes, the attribution of the CO₂ concentration to the back-trajectories after 2000 is made depending on the lat/lon position of the air parcel for a given time. We only averaged the 3-hourly CO₂ mole fraction from CarbonTracker in daily CO₂ concentration that we use to initiate the trajectories.

Page 10-11: Are the particles in each bin spread evenly throughout? I.e. is their initial location in the center of the bin or evenly distributed in time and space? this may have been mentioned earlier but should be clarified here. (Page 11, line 6 discusses how many particles there are per bin, so this would be a good place to discuss how they are distributed in time over the month and in space).

The air parcels are distributed uniformly in longitude $2^\circ/\cos(\text{lat})$ and every 2° for the latitude. At the end of each month, we distribute at each altitude level 10250 particles which are driven with the ERA-Interim winds.

Page 12, L10-15, this discussion of diffusion answers my previous question about the dif-

ference in the models - it could be briefly mentioned earlier in the paper to clarify this.

Okay good! In this section, we calculate these reconstructions at a given location and therefore need to proceed differently than the global reconstructions.

Page 13-14. Perhaps you could clarify why this 6-month dispersion of particles is required before the particles are tagged with the value of the global reconstruction. I would think you could just sample your global distribution at the location of the observations?

As the local variability sampled by in situ measurements is mainly due to the chaotic properties of advection, it is necessary to integrate backward in time to reconstruct this variability from the grossly sampled global reconstruction. Launching only one parcel from each measurement point would reconstruct fluctuations growing with the backward integration time. Using a cloud of diffusive trajectories generates a reconstruction that does not depend on this integration time once it is long enough (see Legras et al., 2005).

Page 14, Line 21, reference to figure needed in text here.

Corrected

Page 14, L23: This is the first reference to a CI for the reconstruction. How is it calculated/obtained? The CI or uncertainty is very important to anyone who would be using this product, and the methods for its calculation should be well documented in this paper. Presumably some of this comes from the particle/diffusion release of the 6-month trajectories from the flight, but there should be an uncertainty associated with the global CO₂ product as well.

The confidence interval of the reconstructions at a given location is clearly an upper bound of the uncertainty of the global product.

Fig. 2: I would like to see differences in addition to the time series of mole fractions, or correlation plots with R² statistics related to bias and/or RMSE; i.e. a more quantitative comparison. A plot or a table of statistics, or perhaps a single plot with all the flights colored differently, would not add too much length or bulk to the paper.

Fig. 3 shows the differences well enough that it seems fine, and statistics would be difficult to calculate for such a small sample.

The correlation and R^2 statistics have been added in a separate figure.

Page 15, L23: We have no information on the error calculation so this sentence is confusing.

This paragraph has been rewritten.

Page 16, L 13: CONTRAIL should be described somewhere (in the section where SOLVE and the balloon flights are discussed in Section 3.2).

A paragraph describing CONTRAIL data have been added

Page 16, section 5.3, why is CONTRAIL treated differently from Solve and the balloon observations, and no 6-month diffusion of particles is conducted? (this is the same as my earlier question about why that was done for those observations, there should be some explanation for this).

The SOLVE and balloon observations are very localized compare to the CONTRAIL that has a large spatial and vertical coverage over several years as shown in Sawa et al., 2008. SOLVE and balloon observations were used to evaluate the ability to reproduce high resolution in situ measurements while CONTRAIL is a more direct test of the global distribution.

Fig 4, some uncertainty bounds should be calculated and shown for the global initialisation using TRACZILLA. (could be based on differences from Contrail, or based on the spread of values of the trajectories that constructed each bin). Some error is likely also introduced by the choice of the level in CarbonTracker (looking at any vertical gradients in CT product could help quantify this, perhaps it is very small).

We have calculated the standard deviation using the spread of values of the trajectories that constructed each bin.

Page 16: Why was 15 days chosen, was it because that delay fit the data best?

Yes! 15 days was chosen because that delay fit the data best at this height but we recovered the 2 months delay at 19-18km above the tropopause.

Page 16, last line: give mean, standard deviation of differences?

Page 17, Line 6-7: in this period, discussed further in section xX.

Page 17, Line 18: awkward description of the biospheric CO₂ seasonal cycle

Page 17, L19: CO₂ concentration in the UT increases (subject /verb agreement in this sentence).

Fig 5, caption should indicate that the source of this data is the CO₂ global reconstruction

Done

Page 18:

L3: tropics

L13 maximum

L14 observations, also refer to CONTRAIL here to clarify this is the same data set used in the earlier comparisons

L25 isolates

L26 no comma after effect, also this is still awkward phrasing

L27 homogenizes

L28 containment

corrected and sentences rephrased

Page 19:

L9 profiles.... exhibit

L10 I dont think interspelled is a word, please rephrase (perhaps interspersed, or alternating?).

L16, processus (process)

L16, injects

corrected and sentences rephrased

Page 20

L7 that should be which

L11 "This good agreement demonstrates that the Lagrangian model (TRACZILLA) ..."

L23-25: awk sentences (high horizontal mixed and uniformise)

L17 extent

L29 troposphere

corrected and sentences rephrased

Page 21

L1 variability

L1 subject/verb (variability are)

L11 measurements should not be capitalized

corrected and sentences rephrased

Answer to anonymous referee #3

We thank referee #3 for his comments and suggestions. Comments by the referee are highlighted and followed by our answers.

Major Comments:

1. *My only major point is the quality of the text (outline, titles of the sections and sub sections). I am not a native speaker but my feeling is that not only the structure of the paper but also the quality of many sentences can be significantly improved.*

We have rephrase some of the titles and improve the discussions

Minor Comments:

1. *Abstract L 1*

please remove relevant. You can only hope that this will be a relevant data set.

Done

2. *Abstract L 8*

I would replace guided by driven

Done

3. *Abstract L 11*

...with mid-latitude vertical profiles measured in situ from aircraft and balloons exhibit a remarkable agreement... (you should not give a complete description of the used data in the abstract)

Done

4. *Abstract L 17*

...out of the tropics to the mid and high latitude stratosphere (but mainly into the northern lowermost stratosphere around 15 km - is it not something that follows from your Fig. 6)

Yes! Rephrase

5. *Abstract L 21*

...and is nearly constant above 35 km (is it what you would like to say ?)

Yes exact! Rephrase

6. *P.4 L3*

These studies.... - please rewrite this sentence

Rephrase

7. *P.4 L 12*

...to help to validate the stratospheric representation in global CTMs...

Rephrase

8. *P.4 L14*

...to the very localized in situ observations which have high spatial resolution, a large spatial...

Rephrase

9. *P.4 L18*

Chardin et al showed.... (please rewrite this sentence)

Rephrase

10. **P.5 L5**

...are also weak... (what do you mean ?)

Rephrase

11. **P.5 L21/22**

models lack of realistic stratospheric influence - not clear, please explain

Rephrase

12. **P.6 L10**

The small scale variability....and the scarcity of suitable observations ... (I would recommend to reformulate this sentence)

Rephrase

13. **P.7 L14 Trajectory starting....(you are using backward trajectories, so maybe you would like to write: Trajectories reaching the boundary layer during the backward integration...)**

The reconstruct with model is done in two steps:

1-) We calculate the trajectories by integrating them backward in time.

2-) We post-treat these trajectories by assigning them within CarbonTracker CO2 in the troposphere at 500hPa.

14. **Section: Data**

You should shortly describe here the aim of both upcoming subsections

Done

15. **Section 4 and 4.1**

I think, you should use a different title like Reconstruction of CO2 and describe it accordingly. Initialization is a very misleading term. So you use backward trajectories

plus Carbon Tracker/WDCGG data in the boundary layer to reconstruct CO2 everywhere in the UT/S region. Please reformulate the text between L10 and L20....

Title and sentences rephrased

16. **P.11 L14**

Maybe you should change b_0 (which is too close to b) to something different.

Done

17. **P.12 eq (2) and (3)**

Maybe you should avoid to introduce Corr, i.e. use only one formula in two lines

We have used one line formula as suggested.

18. **P.11 eq. (1)**

You also did not clearly explain that you need your eq. (1)-(3) only for CO2 reconstruction cases which are older than 10 years and you do not have any information from the backward trajectories. Maybe you should reformulate some sentences...

Yes right. We include these comments explicitly.

19. **P.12 4.2.1**

Once again: for me it not a flight track initialization but much more a Lagrangian reconstruction of CO2 along the flight track - maybe you should reformulate it

Yes! Rephrase

20. **P.12-13**

You repeat here many arguments and formulations from Legras et al., 2005. I do not think this is necessary. I strongly recommend to remove this part and cite the original papers. Instead of this, you should better explain your eq (7), i.e. how displacement in the geometric space is related to a displacement in θ -space.

We have shortened this discussion.

21. **P.14, 4.2.2**

Once again, please use the term Reconstruction in the title of this section and change accordingly the following text...

Done!

22. **P.14 5.1**

Here, I miss any reference to Fig 2.

Corrected

23. **P.14 L 24**

Your abbreviations of the dates are not clear and maybe you should the notation like 26th February 2000, etc.

Figure notation is included in order to overcome this.

24. **P.15 L 6**

aging vortex core

Corrected

25. **P.15 L 7**

corrective step in this instance - I do not understand what you mean

This was belonging to the applied correction to the remaining air parcels into the stratosphere after the 10-yr backward integration using equations 1-3.

26. **P.16 L 6**

accurate - please remove it

Done

27. **P.16 L 8**

...from the global reconstruction calculated by...

Rephrase

28. *In Fig 4a and 4b you denote the regions as tropospheric and stratospheric boundaries. However, in the model there is only one lower boundary prescribed by the Carbon Tracker values. You should exactly say what you mean with your boundaries. For me, 4a validate your model in the middle troposphere around 7-9 km and 4b in the region between 16-17 km. Please clarify this point.*

What we mean here by tropospheric boundaries is that we evaluate our model in the middle troposphere around 7-9 km which is not far away our trajectories CO₂ assignation and by stratospheric boundaries the near tropical tropopause layer where the air enters the stratosphere. We have improved it.

29. *P.17 L 3*

with respect to

Done.

30. *P.17 L 4-7*

please give a more detailed explanation of the discrepancies. as it will be confirmed shortly could be replaced by We discuss this point in the next section.

Need to be done

31. *P.17 L 12*

...derived from our Lagrangian reconstruction

Done

32. *P.17 L 19*

grows

Done

33. *P.17 L 21*

in the southern hemisphere

Done

34. **P.18 L 4**
propagates

Done

35. **P.18 L 5**
is removed from the atmosphere due to...

Done

36. **P.18 L 6**
into the lower stratosphere driven by the lower branch...

Done

37. **P.18 L 9**
west side (can you give more detail how the Asian monsoon anticyclone contributes to this transport

We notice just the Asian monsoon anticyclone contributes to this transport as well discussed in Park et al., 2009, Randel et al. 2010 and Pan et al., 2016 which were specifically focused on this contribution on pollutant transport. In our fig.6(a,b), we can clearly see the "cheminay" further during the summer months June and July between 10-40° N.

38. **P.18 L 13**
which is maximum...

Done

39. **P.19 L 1-2 localized gradient, etc. - please give a more detailed explanation**

Done

Answer to anonymous referee #4

We thank referee #4 for his comments and suggestions. Comments by the referee are highlighted and followed by our answers.

Major Comments:

1. *The evaluation of the data set is very specific and done on the basis of very limited and selected case studies: SOLVE only covers the high Arctic in winter 1999/2000. Similarly the four profiles are arbitrary snap shots, showing disagreement in one case, which is not even tried to be explained sufficiently (p.15).*

The SOLVE dataset is not limited to the polar regions as we have included test flights and transit flights in the extratropics and the mid-latitude, that is six flights out of 12. Moreover, testing CO₂ reconstructions in the polar region during winter is, a priori, the most difficult situation since the polar air is old and should cumulate all the errors in the transport. Regarding balloons, the number of flights with high quality CO₂ data is unfortunately very small and only four are available in the 2000-2010 period. We have also extensively used CONTRAIL data which are partly stratospheric at mid and high latitude over a range of 4 years. Therefore we did our best to evaluate our dataset with the available observations. The sparseness of CO₂ in the stratosphere is one of the reason of our work. See also our answer to item 3 below regarding the comparison with estimates of the age of air.

2. *Further the discussion of the data and differences to literature is very superficial: The time delay between the cycles is mentioned but not appropriately discussed, since Boering et al., 1996 use N₂O = 310 ppbv as reference values to determine CO₂ at the tropopause, whereas in the manuscript a fixed altitude level is used, which is at or even below the tropical tropopause. This is not considered in the manuscript. Instead the authors conclude "...We recover a two-month delay at higher altitude in the layer 18–19km (not shown). The origin of this discrepancy is unclear but is perhaps due to the*

fact that previous studies merge measurements in the deep tropics and the subtropics.”

There is actually no contradiction with Boering et al., 1996, as our 15 days delay is valid in a layer under the tropopause where the air is renewed by convection while the estimate of Boering et al. is associated with the lower tropical stratosphere where the age of air with respect to the crossing of the tropopause grows rapidly with altitude. It is about 6 months at 20 km (see Diallo et al., 2012, where the prediction of our trajectories is shown to fit the ER-2 measurements) and it is therefore not surprising that we find a delay of a 2 months in the 18–19 km layer in agreement with Boering et al. We have modified the discussion.

3. *This is not satisfying for a quantitative reference data set over 10 years. Given the statements in the abstract (l.3-6: "...This product can be used for model and satellite validation in the UT/S, as a prior for inversion modelling and mainly to analyse a plausible feature of the stratospheric-tropospheric exchange as well as the stratospheric circulation and its variability..."), a quantitative assessment and evaluation also of the stratospheric data is required, e.g. by using age of air diagnostics. This would allow for comparison with other data sources or diagnostics (e.g Eyring et al., 2006; Haenel et al., 2015). For this the authors could also include SF6 to their analysis, since the authors conclude that their results hold for any long-lived species (p.20, l.14). Even if SF6 cannot be directly included, the age of air information can be inferred from the data. This would provide a quantitative comparison to evaluate the results on the basis of CO2 and the consistency of the results within the model. It would further help to evaluate their stratospheric data using satellite observations of e.g. MIPAS SF6 in regions where the in-situ data are sparse or absent. Even without SF6 the calculation of age of air allows for comparison with other data sets.*

The same trajectories have been used in Diallo et al. (2012) to calculate the mean age of stratospheric air which has been compared with age estimates based on CO₂ but also SF₆ and N₂O measurements, and also with the GEOSCCM model. It should be observed that SF₆ is photolysed at high altitude and therefore cannot be as easily interpreted as CO₂,

especially at high latitude during winter where a large amount of air has descended from high altitude [more detail see Stiller et al., 2008, 2012].

Minor Comments:

1. *Abstract: Last sentence: Please clarify the sentence and specify: theres a contradiction: decrease or constant? Decrease of CO2 to 35 km or constant, constant with altitude above?*

Done

2. *Introduction: Do you need the first sentence?*

Removed

3. *p.3, l.2-4: The increase of green house gases does not increase tropical upwelling mass flux, it is the effect on atmospheric temperature structure and wave propagation.*

Right. Corrected.

4. *p.3, l.6: stratosphere instead of atmosphere? CO2 is destroyed in the upper atmosphere.*

Done

5. *p.4, l.7: Here you need to mention the Engel et al., (2009) study - not on the previous page (l.25), since it is not based on airborne measurements.*

Done

6. *p.7, l.2: You forgot TRACZILLA*

Done

7. *p.8, l.20-22 (and l.10 ff): What does this mean: Similarly to 1989-1999.... only at 5 km ? Please specify the altitude criterion of the selected stations for 1989-1999. How many stations contribute? It would be good to show a 3D distribution of the boundary condition: e.g. a zonal mean plot with latitude as y axis, time as x-axis and CO2 as*

iso surface above to see the global distribution and allow for comparison with e.g. the NOAA CCGG data.

Corrected. The 1989-1999 is based on ground stations that are far from sources. The criterion to select the ground stations is that the elevation is high enough to neglect the variability due to localized sources at ground level. The ground boundary condition would project on a 2D map but it contains a large amount of variability which is not relevant here. We choose the 500 hPa surface as a boundary condition after 2000 in order to filter out the surface fluctuation and reduce the number of needed trajectories. The CarbonTracker CO₂ which is used to initiate our trajectories is perhaps the best currently available tropospheric CO₂ in the range 2000–2014 because it assimilates all available observations, including CCGG data, to produce a 3D distribution of CO₂ in the troposphere. Therefore a comparison with CCGG data will be redundant and anyway we do not aim at outperforming CarbonTracker in its domain of validity. Our dataset applies above 500 hPa only. The NOAA CCGG webpage highlights the Carbontracker CO₂ <http://www.esrl.noaa.gov/gmd/ccgg/>.

8. *p.10, l.11: Wrong sentence? Something is missing...?*

corrected

9. *p.11, l.3: Is a zonal mean calculated? What are typical numbers of parcels per box?*

We launch 10.255 particles per levels and we have 30 levels. The discretization is described in section 2. longitude=2/cos(lat) and latitude=2 degrees. In the tropics we have 180 particles that decrease as 2/cos(lat).

10. *p.12, l.4: Chapter title: The term validation is used, but one cant validate the results, since you can have agreement for the wrong reasons. Therefore I suggest the term evaluation.*

Okay... Even if here it's not the case.

11. *p.12, l.15 ff.: The exponential factor b clearly depends on the driving data set. Does the exponential factor b further depend on the choice of the trajectory model and needs in principle to be determined for each individual trajectory model?*

The exponential factor is a statistical quantity which can only be defined for a large ensemble of trajectories and not for a single one. It is defined for the whole stratosphere at a given date but varies very little. We made trials of defining by averaging in time for a given set of latitudes and altitudes with negligible effects.

12. *p.13, eqn.4/5: Please clarify the notation of vectors, scalar products and scalar quantities. Why is the **t** in bold font?*

Corrected.

13. *p.13, l.11: "... kappa defined by the user..." Is k (kappa) chosen to have the same value in the whole atmosphere? Please add a word, which values have been selected or how a user has to define Kappa.*

Yes the diffusion is chosen to be the same for the whole atmosphere. We have replaced this discussion by a reference to Legras et al., 2005 where a thorough discussion of the determination of diffusion is provided. It should be recalled that diffusive dispersion of trajectories is only effective during the first 3 or 4 days of the backward integration, hence the chosen diffusivity has to be valid in the lower stratosphere only.

14. *p.15, l.19-23: Please explain, how a cold front (which is a tropospheric feature) can affect the CO2 at 25 km altitude. The inset in Fig.3. is too small and does not contain a legend. It is further unclear, why the PV gradient should be associated with a CO2 gradient. This paragraph sounds very weird or almost wrong.*

There is nothing weird here but the mere usage of common concepts in dynamic meteorology. When submitted to chaotic stirring, all long-lived tracers that are bound to preserve their tracer-tracer relation tend to align their contours and therefore the CO₂ contours are likely to follow the PV contours and to exhibit high gradients at the same location of high PV gradients (as would ozone or N₂O do as well). A cold front is the surface signature of a deep structure that penetrates the stratosphere. The map shows the PV distribution at 18.5 km well above the mid-latitude tropopause located at about 11 km. A similar case has been studied in depth by Pisso and Legras, 2008. See also Miyazaki et al., 2009.

15. *p.15, l.27-29: "... The mean in situ CO₂ from observations is much more spread in the high latitude profile (44° N) above 15 km. There is not a clear explanation about these observed fluctuations on the in situ CO₂-profile.":* What do the authors want to say with such a statement? What does this mean for the comparison? What does it mean for a data set, which is intended to serve as a reference for model evaluation from 2000-2010, if one out of four stratospheric profiles does not fit the observations?

It is well known that modelling tracer distribution is highly prone to transport errors in the region of large gradients. It was shown by Pisso and Legras, 2008, that the global pattern is weakly affected but a small displacement can in this case generate important deviations. Therefore it is expected that the test will fail in such a case which cannot be used as a reference. The discussion has been rewritten.

16. *p.19, l.20 and l.23: What is meant with "the subtropical barrier" in this paper? Do you mean the subtropical jet at the tropopause, which exhibits a seasonality with weaker PV gradients and high permeability in summer? The subtropical barrier normally denotes the boundary of the (leaky) tropical pipe in the overworld (e.g. Palazzi et al., 2009), which does not show the same variability as the STJ and has a different generating mechanism.*

By subtropical barrier we mean "the subtropical jet at the tropopause, which exhibits a seasonality with weaker PV gradients and high permeability in summer". This is a common terminology in dynamic meteorology where it has been known for a long time that the jet centered at 200 hPa and about 30°N and S inhibits exchanges between the upper tropical troposphere and the lowermost extra-tropical stratosphere.

17. *Fig.2: Please include potential temperature along the flight track as additional information. Otherwise the information on the plots is without any relevance for a scientific interpretation and an estimate of the quality of the model capabilities. Does the gray area refer to the variability of the data in a bin? If not, how is the error calculated?*

Potential temperature along the flight-track has been included. The gray area is 95% confidence interval estimated from the model results.

18. *Fig.5: The continuous color bar is not consistent with the figures, which have discrete colours. Please provide a discrete color bar legend.*

Done.