Responses to referees’ comments (ACPD-13-26327-2013)

We thank anonymous referees for their thoughtful and constructive comments and helpful suggestions. We also thank Dr. Gerbig for his careful handling of this manuscript.

We have fully considered the referees’ comments in the revision and improved the manuscript accordingly. The referees’ questions are italicized and are immediately followed by our answers.

Anonymous Referee #1

The manuscript provides a well-written comprehensive study on regional sources and sinks of column-averaged atmospheric CO2 mole fraction (XCO2) derived from GOSAT observations. Three different inversions were run and compared to independent data from aircraft (HIPPO) and ground-based total column XCO2 measurements (TCCON). The text is somewhat lengthy, which is largely justified by a detailed and comprehensive introduction and description of the inversion system and used data sets.

Answer: We thank Anonymous Referee #1 for the positive and constructive comments, which we used to improve our manuscript. More detailed description of the changes we made per comment is given below.

My main concern is the fact that the three different inversions produce very different results for the net fluxes of individual regions. For example, depending on the choice of inversion, North America could either be a strong source or a weak sink – even though all inversions agree well with the independent observation data. Though the authors try to explain why one inversion produces one or the other result, this leaves many questions. First of all, the reasoning behind setting up the three inversions with different parameters is not explained well. Therefore, it is hard to judge which of the three setups should be closest to the real world. This leaves the impression that the choice of inversion parameters is arbitrary and there is no optimal choice. If that is so, the authors would be overinterpreting their results. In that case, the system would simply not be able to provide useful regional estimates - at least not on the level of the TransCom regions that were used here. One could imagine, that a more systematic approach *like an ensemble run could shine some light on this issue. However, this is not exactly my field of expertise.*

Answer: As we noted on page 4, lines 10-11, the goal of the paper was "an investigation of the impact of biases in the GOSAT CO_2 data on regional flux estimates of CO_2." The basic setting of three inversions was identical. The differences among the three inversions were due to the three different subsets of the data, reflecting different treatment of the biases. We agree with the reviewer, that our inversion results suggest that the regional flux estimates are not reliable. We stated this on page 28, line 9-10, "although the global flux estimates inferred in the different inversion analyses presented here were robust, the regional flux estimates were less reliable." We chose the scale of the TransCom regions to aggregate the fluxes, but as we noted in the paragraph (page 28, line 14-17) "An objective approach is clearly needed to determine the minimum spatial scales at which the fluxes can be reliably quantified. There is also a critical need for additional independent data to better evaluate the inferred fluxes." We agree with the reviewer that this is a cause for concern. As pointed out in our introduction section, GOSAT is the first successfully-launched satellite designed to monitor total atmospheric column carbon dioxide (CO_2) globally from space. Retrieval of XCO_2 data from the satellite signal is a challenging task, and fortunately the XCO_2 data quality has continued to improve. For example, the systematic bias has been improved from larger than 5 ppm for ACOS version 2.8 to less than 1 ppm for version 2.10. Our work showed that the data biases through version 2.10 were still too large to obtain reliable regional flux estimates. We anticipate that newer versions of the XCO_2 data product will be even better and, hopefully, will provide more reliable regional flux estimates.

General remarks:

- use past tense consistently in the manuscript for all the descriptions of what you did as well as your findings.

Answer: We did our best to change the tense, where it did not make the discussion cumbersome.

- labels are far too small in most plots!

Answer: We enlarged the font size for most plots.

- use consistent color code for all figures

Answer: We modified the color code for all figures to make them consistent.

- display latitudes with N/S and longitudes with E/W instead of +/-. Especially for longitudes, there is no "natural" sign convention.

Answer: Corrected.

- I would prefer if all time axes were either month or day of year

Answer: We changed all time axes to month.

Individual corrections:

- p. 26334: the information about the XCO2_A, XCO2_B, and XCO2_C datasets is spread over several paragraphs. You should have a table that summarizes the properties of the three data sets as well as key parameters of runs A-C.

Answer: We described the GOSAT data (XCO2_A, XCO2_B, and XCO2_C) and the relevant filter and bias correction information in one paragraph, the last paragraph in Section in 2.1.1. Discussion of the data uncertainties and adjustment of the initial conditions, etc..., for each dataset were described in other sections of the manuscript since they were relevant for the description of the assimilation configuration. We cannot integrate this information into one table.

- p. 26335, l. 17: "... is then is computed ..."

Answer: Corrected.

- s. 3.2.1: all biases should be reported as bias +/- std deviation.

Answer: Done.

- p. 26348: mark the stations you used on the map in Fig. 2. Alternatively, provide a table that lists all used station names and locations.

Answer: We used data from all 78 flask stations and 13 TCCON sites shown in Figure 2. We now specifically give the lat/lon information for each station that is presented in Tables 2 and 3.

- p. 26349: again, always provide biases along with std deviation

Answer: Here is a little difference from sect. 3.2.1 mentioned above, we have specifically used 'mean' bias or difference, so we did not attach standard deviation.

Figures:

- Figure 1: axis labels very small, explain choice of latitude bands

Answer: Labels were enlarged. The latitude bands were chosen to separate the tropics/subtropics (0-30) from the extratropics (30-65), excluding the Arctic and Antarctic regions; however, the values shown in the plot are the latitude boundaries of the 4×5 grid boxes that encompass the 0-30 and 30-65 zones.

- Figure 2: TCCON diamonds not ideal (use crosses?), labeling unreadable

Answer: We changed to black crosses to indicate TCCON sites, and enlarged labels.

- Figure 3: labels and color scale unreadable

Answer: Labels were enlarged.

- Figure 4: strange - Europe exactly 0 as a priori?

Answer: It is not exactly zero, but very small. As we defined this term, it includes biosphere exchange and biomass burning, excluding fossil fuel emission, so with an annual neutralized biosphere (GPP+RSP=0) plus a very small biomass burning source, the total flux is a number very close to zero for this region in 2010.

- Figure 5: labels unreadable, should keep consistent color scale with Fig. 4

Answer: Labels were enlarged and the color scale is consistent with Fig. 4.

- Figure 6: should probably be "... smallest uncertainty reduction obtained ..."

Answer: Corrected.

- Figure 7: labels unreadable, mean and standard deviation should be more clearly indicated - not just N(..., ...)

Answer: Labels were enlarged, and mean and standard deviation are clearly indicated.

- Figure 8: plot is too busy, labels too small, the site indicator with the lat/lon in brackets is not very intuitive, would be better to mark the sites on the map in Fig. 2, consistent color code?, indicate year

Answer: we enlarged the labels, deleted lat/lon from these plots, and added lat/lon information in Table 2 and 3, and indicated year in the caption. The color code of Fig. 8 is consistent with Fig.9.

- Figure 9: similar problems as Fig. 8, consistent color code, use TCCON station codes instead?, indicate year

Answer: Changes were made as in Fig. 8.

- Figure 10: is hard to see any difference between the 3 plots. If there is one, maybe you should plot this as difference from a priori (since the HIPPO data and a priori are the same on all plots)

Answer: This figure was redrawn as suggested.

- Figure 11: use consistent color code

Answer: Done.

- Figure 13: labels too small

Answer: We enlarged the labels.

- Figure 14: labels too small

Answer: We enlarged the labels.

- Figure 15: labels too small, lines too weak and poor choice of colors, why is the quality of the plots so bad?

Answer: We have made a new version of the figure.

Anonymous Referee #2

This study investigates the use of GOSAT retrieved total column CO2 for estimating global CO2 fluxes using inverse modeling. Such studies are much needed to support the discussion of what can be expected from space borne C-monitoring and what is needed in the context of future missions. The results confirm the sensitivity of inversion-estimated fluxes to systematic uncertainties in the measurements, and the spatio-temporal variability of measurement coverage. Inconsistencies remaining connecting surface data to total column data, which are corrected without investigating the underlying causes. To avoid confusion on this topic some further effort is needed as will be explained below. Overall, the study is carried out well and results are documented in a transparent manner. In my opinion a few important issues remain in connecting results to conclusions. If these are addressed in a satisfactory manner I see no reason to uphold publication in ACP.

Answer: Thanks for your positive remarks to our effort in using ACOS GOSAT XCO_2 data to constrain atmospheric CO_2 inversion. We really appreciate your constructive suggestions. The comments are a great help to improve the manuscript. Below you can find our detailed responses to the comments.

GENERAL COMMENTS

Bias correction and comparisons to TCCON

Given the importance systematic measurement errors, it is important to accurately define and evaluate the correction methods. Some further attention is needed here.

From the description on page 26334 it is not clear whether there was no bias correction applied to the b2.10 data or that they were filter and corrected as in XCO2_B. The next obvious thing to do is evaluate the agreement of the GOSAT optimized model against TCCON (Table 3) to find out whether the corrections did what they were supposed to do. Comparing RUN_A and RUN_B in table 3, it is not obvious at all that the 4 parameter bias correction improved the fit to TCCON. It is true that this comparison is indirect, since the correction is on the GOSAT data – which are not compared to TCCON here. The question remains whether the bias in correction of the data that went into RUN_B did improve the comparison to TCCON (which I suppose should be the case), and if so why no improvement is seen in the comparison between RUN_B and TCCON.

The next question is how the differences in table 3 and 2 relate to the adjustments to the initial conditions that were applied. I assume that Table 2 is for the uncorrected initial conditions (this info is missing in the caption). It is surprising to see that RUN_C leads to a satisfactory small offset to TCCON as well as the surface data, without any correction to the initial field. In the case of RUN_A and RUN_B there is this ~1 ppm adjustment needed to match surface to total column. Some discussion is needed as to why this is. The 1 ppm correction suggests some problem connecting in situ data to total column data, which may be due to the model or TCCON (provided that the GOSAT data are bias corrected to be close to TCCON). The results for RUN_C, however, rather suggest that there may be no inconsistency problem at all.

Answer: As described on page 6, lines 12-13, the ACOS b2.10 data used in RUN_C was filtered and bias-corrected. As regards the b2.9 data, the 4 parameter bias correction (including filtering and bias) did improve the fit to TCCON as shown by Wunch et al. (2011). However, the comparison with TCCON presented in the manuscript does suggest that the bias-corrected b2.10 data provide better overall agreement with TCCON than the bias-corrected b2.9.

In the original text, on page 26333 line 20, we stated that “the filtering and bias correction schemes were refined and updated for version b2.10.” in order to indicate to our reader that this dataset has already been filtered and bias corrected. We have revise this sentence to read “the filtering and bias correction schemes were refined and updated for version b2.10, and the ACOS b2.10 data used in this study have already been filtered and bias corrected”.

NH Extratropical sink and the seasonal cycle

It is concluded that GOSAT data increase the C-sink in the NH extratropics. This is explained by the larger seasonal cycle amplitude in the GOSAT data compared with the prior, as confirmed in the comparisons to TCCON. In my opinion there is no logical connection between the seasonal cycle amplitude and the annual mean flux. There is some loose argumentation (which could be made explicit) that the GOSAT data provide a stronger constraint in the growing season, and so the seasonal cycle amplitude is corrected more strongly in the summer than in the winter, which would impact the annual mean flux. However, all this depends on what happens in winter. Is it true to that fluxes are adjusted less in winter, because of a weaker observational constraint?

Does it affect the comparison to TCCON during winter? In my opinion this conclusion is made too quickly and needs further support.

Answer: The reviewer is correct in that a greater seasonal cycle amplitude does not necessarily mean a larger annual mean carbon sink, but we did not make an inference between the stronger seasonal cycle and the annual mean a posteriori flux estimates. We concluded that the inversion produced, relative to our a priori, a larger seasonal cycle in the NH extratropical fluxes because, as shown in Figure 5, the wintertime a posteriori fluxes were comparable to the a priori, whereas the summertime a posteriori sinks were much larger than the a priori. We believe that our conclusion is valid. It could be that the fluxes are adjusted more strongly in summer than in winter because GOSAT provides a stronger observational constraint in summer. This was noted by Liu et al. [Carbon Monitoring System Flux estimation and attribution (CMS-Flux): Impact of ACOS-GOSAT XCO_2 sampling on the inference of terrestrial biospheric sources and sinks, submitted to Tellus B, 2013]. Regarding the comparison to TCCON, the agreement between the a posteriori CO_2 field and the TCCON data is as good in winter as it is in summer. We have modified the text at the beginning of the conclusions slightly to state that: (Page 26, line 22-24) ‘The inversions significantly increased the uptake estimate in the northern extratropics during the growing season, suggesting that the a priori fluxes may have underestimated the seasonal cycle in the northern extratropics.”

Measurement footprint analysis

The difference between the inferred fluxes over the European and American continents are explained by differences in data coverage and footprints. Although some results are presented supporting the ideas of how these factors influence the inversion derived fluxes, it is not clear how important they really are. E.g there is no prove of the impact of the Eurasian data on the flux uncertainty reduction over America. It might be that the measurements over the American continent are far more important. Without stronger evidence of these relationships the wording in the conclusion section regarding the important of long-range transport and the needed for measurement coverage over the ocean should be more careful. One might think of additional experiments to provide such prove, such as truncating response functions (shortening the optimization time window) or data thinning. This may be out of scope of the paper, but right now the conclusions are not sufficiently supported by the evidence that is provided.

Answer: The “footprint” shown in Figures 13 and 14 are the influence functions that relate changes in atmospheric CO_2 to the fluxes. In the context of a regional inversion analysis these figures are showing the Jacobian for the inversion, so they allow you to make a relative comparison of how the different flux regions will be influenced by the observations. However, it was not our intention to make a definitive statement about the importance of long-transport transport. We have modified the text in the conclusions on page 27, line 14-23. It now reads: “We note that the analysis presented here was an initial attempt at understanding the inversion results in the context of the transit times. It suggests that the North American flux estimates are more strongly influenced by long-range transport and should, therefore, be more sensitive to regionally varying biases in the observations and to model transport errors. In addition, the low sensitivity of the European flux estimates to observations outside of Eurasia could explain why the inferred European flux estimates are more robust across the three different XCO_2 datasets. However, there is clear need for more detailed analyses to better characterize the sensitivity of the inferred regional flux estimates to the transport pathways and transit times associated with long-range transport of the continental CO_2 signals.”

SPECIFIC COMMENTS

Page 26335, line 5: What is the logic of quantifying the representation error by comparing the fit of the a priori model to the data. This difference is accounted for in part by the a priori flux uncertainty, which seems to be counted twice in this approach.

Answer: In the cost function, Equation (6), the a priori flux uncertainty is specified in the Sa matrix. The representativeness errors discussed in Section 2.1.2 are captured in the observation error covariance matrix, So. The representativeness errors represent the inability of the coarse resolution of the model to capture the subgrid-scale variability that the flask measurements (which are point measurements) will capture. In the first term of the cost function, the model-data mismatch $H(\mathbf{x})-\mathbf{y}$, where y are the observations and $H(\mathbf{x})$ is the model simulation of the observations, will reflect the inability of the model to capture this subgrid-scale variability. For flask measurements at the surface, the representativeness errors will be significantly larger than the measurement errors (which are also captured by So), but difficult to quantify. Therefore, we used the statistics of the model-data mismatch to characterize the observation error So. The fundamental assumption that is made in all flux inversion analyses is that the mean model-data mismatch is attributed to discrepancies in the a priori flux estimates and that the random variability around this mean is due to the observation errors, as reflected in So. As we noted in the text, this approach has been used by others, such as Palmer et al., (2003) and Heald et al. (2004).

Page 26337: at what altitude are aircraft emissions introduced in the model?

Answer: The aircraft emissions are a 3-D dataset. As described in Nassar et al. (2010), which is cited in the text, the altitude ranges from surface to about 12 km, with a small peak below 1 km (reflecting takeoff) and a maximum in emissions between 9 to 12 km.

Page 26340: The uncertainties are scaled to satisfy the chi2=1 criterion. It is not clear how this is done for the observation and regularization terms separately. There is only one chi2, which seems to be used to infer 2 scaling factors. Or does 'reduced chi2' mean that the observation and regularization terms should contribute equally to the overall chi2. I don't see a reason why this should be the case.

Answer: You are right that there is only one χ^2, and it cannot be used to infer 2 scaling factors simultaneously. Actually we mainly used χ^2 to adjust the scaling factor for observation uncertainties. However we also set a principle that regularization terms cannot exceed 10% of the overall χ^2, usually around 5% of the overall χ^2. At the beginning of our study, we did some experiments to adjust a priori flux uncertainties according to this principle and our priori knowledge of different emissions. Once prior uncertainty scaling factors were determined, we focused on using χ^2 to adjust the scaling factors for observation uncertainties of various observation datasets.

Page 26340: Some scaling factors of a priori uncertainties are postulated without motivation. What evidence is supporting these corrections?

Answer: As mentioned above it is important to remember that the a priori term in the cost function is just a regularization term. Because of the long assimilation window used in our study, a significant amount of data is ingested so the inversion is strongly driven by the observational term in the cost function. This is typical with global inversions using satellite data. In these inversions it is common to inflate the a priori term to enhance its contribution to the cost function. For example, Muller and Stavrakou (Inversion of CO and NOx emissions using the adjoint of the IMAGES model, Atmos. Chem. Phys., 5, 1157–1186, 2005) uniformly scaled down by a factor of 5 to their specified a priori errors to obtain the desired balance between the observational and a priori terms in the cost function. Here, we started with relatively small uncertainties for the grid level flux estimates and then inflated them to achieve an a priori contribution to the cost function that was about 10%. For example, the estimates of biomass burning emissions are uncertain to about 20% on global, annual scales (van der Werf et al., 2010), and we inflated this uncertainty to 38% to contain the exacerbated uncertainties at the grid level. The prior uncertainties of terrestrial ecosystem flux components are based on the prior knowledge that the GPP estimates are uncertain to ~6%-13% on global, annual scales (Chen et al., 2012), and we inflated these uncertainties to 22% to contain larger uncertainties at grid level. For ocean fluxes, we assumed an a priori uncertainty of 44%, to keep the relative proportions for land and ocean in the range of those used in previous studies (Deng and Chen, 2011).

We have modified the text to give the reader the appropriate reference for our assumed a priori errors. The paragraph now states: (page 11, line 21 to page 12, line 2)"…According to Marland et al. (2008), the uncertainty for estimates of global fossil fuel emissions is about 6%. However, in constructing Sa, we assigned 16% for the uncertainty of the fossil fuel emissions in each month and each model grid box. For biomass burning, we started with an assumed uncertainty of 20%, a global annual uncertainty estimate (van der Werf et al., 2010), that was then inflated to 38% for emissions in each month and in each model grid box. The annual GPP estimate for 2010 is -119.5 Pg C and we assigned an uncertainty of 22% of the GPP estimates in each 3hr time step and in each model grid based on global annual uncertainty estimates of 10% to 13% (Chen et al., 2012). TER, which is the sum of autotrophic and heterotrophic respiration, was specified to be 119.5 Pg C in 2010 since we assumed an annual balanced biosphere. We also assigned 22% of the prior estimates in each 3hr time step and in each model grid as the prior TER uncertainty. For the ocean flux we assumed an a priori uncertainty of 44%, to keep the relative proportions for land and ocean in the range of those used in previous studies (Deng and Chen, 2011).”

Page 26343: It is suggested that 2 inversions do not satisfy the global budget constraint as derived from the global growth rate. Does it mean that these inversions don’t reproduce the observed CO2 increase? Or that something is going wrong internally?

Answer: As far as we know, the global growth rate is derived from global mean atmospheric mixing ratio (concentration) at the beginning and the end of a year. But how to calculate the global mean (whole atmosphere) mixing ratio (concentration) and how many observations should be used to calculate these global mean mixing ratios are not so easy to be accurately determined. In fact, atmospheric inversions also provide a means for deriving the global growth rate using an atmospheric transport model and atmospheric CO2 observations. The current estimates of the global annual a posteriori uncertainties are of ~±0.5 PgC each year, and this is much larger than the ±0.13 Pg C error estimate for the 2010 global growth rate. Though two of the inversions do not satisfy the global budget constraint as derived from the global growth rate, it is not easy to determine whether these two inversions did or did not reproduce the real global growth rate of 2010 within acceptable errors. It is also important to keep in mind that the inferred growth rate in the inversion will be sensitive to the initial conditions imposed at the start of the inversion. As discussed in Section 2.5, we removed the global mean bias in the initial CO2 fields at the start of the inversion, but residual regional biases and the non-uniform spatial-temporal sampling of the GOSAT observations will impact the inferred flux estimates even after half a year’s spin-up period.

Page 26346: The effect of neglecting spatial correlations is not so easy to assess, since it will depend on the scale at which fluxes are evaluated. Local fluxes may become better constrained by adding positive correlation to regions that are better constrained by the measurements.

Answer: We agree with the reviewer. We have changed the wording to “Neglect of spatial and temporal correlations in the a priori error covariance matrix would also result in discrepancies in the predicted a posteriori errors and, consequently, in errors in the estimated uncertainty reduction.”

Page 26347, line 24: “CO2 concentrations” io “CO2 fluxes”

Answer: The reviewer is correct, that we are comparing the a posteriori concentrations, but the objective is to evaluate the fluxes. Unfortunately, due to the lack of CO2 flux measurements that are representative of the scale of the model resolution, we are forced to indirectly evaluate the fluxes using the a posteriori CO2 concentrations. We prefer to leave the text unchanged to make the point clear that our goal is indeed evaluation of the fluxes.

Page 26348, line 19: Is sounds kind of strange to choose poor priors intentionally. Then you end up intentionally degrading the quality of your solution also, which doesn’t make sense.

Answer: It does sound odd and the reviewer is correct in noting that we will recover the poor prior where the observations offer little information on the flux estimates. However, since we cannot, at this time, quantify the contribution of the prior to the posterior flux estimates, this approach enables us to more easily evaluate whether the information in the data is sufficient to constrain the flux estimates. If we started with a very good prior and obtained only a minor adjustment to the fluxes it would be difficult to tell whether that was due to the lack of information in the observations or due to the fact that the GOSAT data are consistent with the prior. Since the GOSAT data are relatively new, we felt that our approach was a valid first step for assessing the utility of the data for quantifying regional flux estimates. We have changed the wording to ‘We started with a neutralized annual a priori flux to better assess the ability of the observations to constrain the flux estimates’.

Table 2: The caption should mention how the initial condition is treated to arrive at this comparison.

Answer: Done. We added ‘These a posteriori simulations use the original initial field that differ from those used for the inverse modeling in RUN_A and RUN_B.’ to the captions of Table 2, Table 3, and Table 4.

Figure 3: It would be useful to be able to compare the differences with a posteriori uncertainty ranges.

Answer: The reviewer is correct, but we cannot estimate the annual posteriori uncertainties directly from the inversion. Annual posteriori uncertainties could be calculated based on the a posteriori uncertainties of monthly fluxes. In our case we have optimized 198720 surface fluxes for 12 months of 2010 (based on the 4°×5° resolution and various CO_2 fluxes in the state vectors), so the a posteriori uncertainty matrix for these monthly fluxes can be as large as 198720 × 198720 (=39489638400), which is approximately 150 terabyte (TB) in 32 bit floating-point type. We are unable to handle this matrix to get these annual uncertainties. Currently we can only report a posteriori uncertainties of monthly fluxes, so we added a sentence at the end of Section 2.4, on page 13, line 3-5, to clarify this issue: “We estimated here the uncertainty on the monthly flux estimates. The large computer memory needed for this approach prohibited us from applying it to estimate annual uncertainties for the regional and global flux estimates.”

Figure 12: Do I understand correctly that the black boxes are also the areas from which CO2 is emitted to obtain the results of Figure 13 and 14 (if so this should be mentioned explicitly). In that case I would think it would be fairer to choose regions of approximately the same size. Otherwise the 1 PgC pulse doesn’t lead to the same emission per m2.

Answer: No, the black boxes indicate the receptor areas over which the transit times are averaged in Figure 15. The CO_2 is emitted according to the spatial distribution of the CO_2 fluxes shown in Figure 12. The spatial pattern represents the actual distribution of the CO_2 fluxes. The total flux from each continental region (North America, Europe, and Asia), however, is scaled to a total flux of 1 Pg C. We have modified the caption for Figure 12 with the following text: “The pulses are produced by emitting CO_2 based on this spatial pattern. The black boxes represent the receptor regions over which the transported pulses are averaged for the transit time analysis shown in Figure 15.”

Figure 13 and 14 are difficult to integrate spatially by eye. As a result, a more intense pulse may show up more prominently than a more diffuse one.

Answer: We have made new plots of Figures 13 and 14 so that they will be more legible.

Figure 13 and 14: The color scales should be the same to allow proper intercomparison.

Answer: We have used the same color scales for the new plots.

Figure 14: The middle panel looks strange to me. It suggests that the pulse takes a month to travel from Europe to Siberia. I suppose this has to do with the transition across the month. Emissions in the final day of April, will contribute to the response on the first of May. If you want to assess the impact on a monthly time scale it would be better to release instantaneous pulses and evaluate after a month. What you get now represents a range of time scales from 1 day to a month.

Answer: One cannot use Figure 14 to infer transit times. To do that, you need to release instantaneous pulses, as suggested by the reviewer. This is what was done for Figure 15 – we emitted the pulses daily. As shown in Figure 15, the European pulse shows up in the Siberian middle troposphere in less than about 1 week. However, the pulse is broad, suggesting that the residence time is a couple of weeks.

Figure 15: If you do instantaneous pulses, then you need to average many of them. Otherwise the results may well be influenced by specific meteorological conditions, whereas the results are interpreted as a more general finding. Else, please improve the quality of this figure (thin lines, poor resolution).

Answer: The reviewer is correct. The pulses are influenced by the meteorological conditions. As we noted in the text "we also examined the transit times for tracers released in 10-day intervals in June and found that the changing synoptic conditions in June did not significantly change the distribution." We have modified Figure 15 and now show the average of the pulses released on June 1st, 10th, and 20th. As we mentioned previously, the distributions are similar.

We have modified the text to (page 25, line 19-23) "Fig 15 shows the transit times to the middle troposphere over the receptor regions in Fig 12 for June 2010 conditions. To reduce the influence of the changing synoptic conditions in June on the distribution of the transit times, we averaged the distribution of transit times obtained from pulse releases of the tracers on June 1st, 10th, and 20th, 2010."