Response to reviewer one: "Elevated uptake of CO2 over Europe inferred from GOSAT XCO2 retrievals: a real phenomenon or an artefact of the analysis?"

We'd like to thank the reviewer's detailed comments, which indeed have helped us improve our manuscript. Below we address all comments (marked by italics) raised by the reviewer.

General comments:

"My main problem with the paper is the following: In the title a question is asked and the abstract suggests that this question is addressed and answered, although the answer remains a bit unclear. Nevertheless, most readers will very likely conclude from the abstract that the answer is: Yes, this is an artefact (e.g., Abstract, page 1991, lines 9- 10: "We show this elevated uptake over Europe could largely be explained by mis-fitting data due to regional biases"). The question is: What exactly is meant by the mentioned artefact? Here the authors refer primarily to one recent peer-reviewed publication and the Feng et al. paper casts doubt on the findings of that publication. The question in the title is related to a key conclusion from the recent paper "Satellite-inferred European carbon sink larger than expected" of Reuter et al., ACP (2014) (see Feng et al., page 1993, lines 8-11, see also below). Feng et al. are essentially aiming at addressing the question if the results shown in Reuter, et al. (2014) are "real or an artefact of the analysis". To investigate this question is important but shown below the analysis performed by Feng et al. is not appropriate to validate or invalidate the results presented in Reuter et al. (2014)"

The title of the paper was chosen intentionally to highlight that we had not reached a robust conclusion about whether the elevated European uptake reported by previous researches was a true phenomenon or an artefact caused by uncharacterized systematic bias of the data or associated with issues of inversion approach. As noted by the reviewer many previous studies, including ours, reported a large European carbon uptake during summer inferred from GOSAT data (e.g., Basu et al., 2014, Deng et al., 2014, Frederic et al., 2014), but Reuter et al (2014) was the first one is to highlight this result as a real phenomenon. Reuter et al might well be correct that this is a real phenomenon and Europe is indeed taking up more carbon than previously thought, but we have proposed an alternative hypothesis involving systematic bias of data within and outside the European region. The artefact in this case is the mis-interpretation of the data if they have uncharacterized biases. Currently there are no sufficient data to rule out that GOSAT X_{CO2} retrievals are indeed biased. Also as shown in revised Appendix A, characterizing and correcting for systematic bias is non-trivial. Mis-characterization tends to weaken observational constraints and compromise a posteriori flux estimates.

Our objective was not to invalidate the work by Reuter et al., but to highlight that other hypotheses are consistent with available data as well. Our intention was to promote further studies in this area. In the revised abstract, we make this point clearer.

Feng et a. write (Abstract lines 20-22): "We find that 50–80% of the elevated European uptake in 2010 inferred from GOSAT data is due to retrievals outside the immediate European region, while most of the remainder can be explained by a sub-ppm retrieval bias over Europe". While these findings are relevant for the flux inversion method used by Feng et al., they do not permit to draw any conclusions with respect to the Reuter et al. findings for a number of reasons, for example, because Reuter et al. are not using any satellite data outside of Europe. Retrieval errors outside of Europe therefore cannot influence the findings of Reuter et al. As clearly explained in Reuter et al. (2014) they only used satellite data over Europe to avoid potential issues with non-European data. The analysis presented by Feng et al. is therefore not appropriate to answer the question posed in the title of the manuscript – at least not with respect to the Reuter et al. publication.

The aim of this paper is to question whether the elevated European uptake inferred from GOSAT data is real or an artefact due to the high sensitivity of flux estimates to small regional biases. We have investigated the relationships between flux estimates and possible biases within and outside European region. Our global (and quasi-regional) inversion experiments shows that a large portion of the enlarged European uptake is related to elevated CO_2 inflow caused by assimilating GOSAT XCO2 retrievals outside the immediate European region. A varying sub-regional bias of 0.5 ppm can explain much of the remaining extra uptake. We do not have a sufficient number of independent observations to preclude the possibility of such regional observation

biases. Also, we show it is non-trivial to develop a robust inversion system to detect and correct these biases. These points have been stressed in the revised manuscript.

In this context it needs to be pointed out that the Reuter et al. (2014) paper is the relevant publication in the context of the Feng et al. paper as can be seen from, e.g., page 1993 lines 8-11: "Consequently, there is an ongoing debate about whether a recent study that shows a large European uptake of CO2 (Reuter et al., 2014) reflects a real phenomenon or is an artefact of uncharacterized regional biases".

See above for the explanation why we highlight Reuter et al. We fully acknowledge other previous studies that report similar results.

Reuter et al. (2014) were not the first who analyzed satellite XCO2 retrievals and found a larger European carbon uptake (see references given in Reuter et al., 2014, in particular Basu et al., 2013, and Chevallier et al., 2014).… The goals of the Reuter et al., 2014, paper are therefore the same as the goals of the Feng et al. manuscript. However, Reuter et al. (2014) present a more appropriate analysis to answer the "real or artefact" question as they specifically address potential issues discussed in the literature which may result in erroneous European carbon fluxes.

We agree that the two papers have similar objectives. The emphasis of our paper is that observation biases can result in an apparent elevated estimate of European uptake, which cannot be ruled out without further dedicated measurements. Similarly, more measurements are needed to invalidate the hypothesis by Reuter et al.

To achieve this, Reuter et al. (2014) used a new CO2 flux inversion method, insensitive to potential (e.g., seasonally varying) retrieval biases, analyzed several satellite XCO2 data products, performed various sensitivity assessments and used satellite data only over Europe to make sure that potential retrieval biases of the satellite data outside of Europe do not adversely impact the European results. […text with no questions…]As explained (see also below), the performed analysis is not appropriate to answer the titled question in general. Any conclusions that can be drawn from the performed analyses are specific to the utilized inversion system. For example, others do not use the EnKF technique,and/or use bias correction schemes, and/or use regional inversions instead of global inversions (see cited literature). The manuscript of Feng et al. shows that small biases have the potential to influence the used inversion system. However this is well known from several previous studies. From this, however the authors conclude that such biases indeed exist and explain the elevated European sink seen by their inversion system. This is not supported by the performed analyses. Their performed analyses are not appropriate to show that such biases are the reason for the observed sink of other inversion systems (especially if large scale or temporal biases are accounted for as done in Reuter et al.

In summary, the title is too general and the abstract and the conclusions should make clear that the manuscript only shows that sub-ppm biases have the potential to explain the elevated sink and that the performed analyses are not appropriate to doubtlessly show that such biases indeed exist and that they are the reason for the observed sink in the used or other inversion systems. The study would be better suited for publication in ACP if more general conclusions were possible.

The title was just chosen to emphasize that there is no consensus about the elevated European uptake of CO2. As stated above as well as in our revised concluding remarks, our work *does not* invalidate the science reported in Reuter et al. The results we reported *are* general: even a small systematic bias can compromise interpretation of observations.

It is difficult to develop a robust way to remove the adverse impacts of uncharacterised observation biases. We have previously applied bias corrections to GOSAT XCH4 inversions (Fraser et al., 2013, 2014), and XCO2 inversions (Feng et al., 2014, IWGGMS). From these studies, we have learned the limitations of the so-called on-line bias correction: 1) unintentionally removing the real signal of the varying surface fluxes, and 2) the method only being effective when the observation biases actually appear as the assumed patterns. An important question about the regional inversions with on-line bias correction is how to specify the bias pattern and its prior uncertainty. As shown in Appendix A, if a priori uncertainty of the bias is too large we

may weaken the observational constraint of the varying monthly mean concentration on the biosphere seasonable cycle across the region, so that we are unable to establish a steady relation between the inferred fluxes and the 'true' regional total. We have detailed this discussion in Appendix A.

Major comments:

1) Abstract, conclusions: The main conclusions are not supported by the performed analysis (see general comments). Additionally, the results of the performed inversion studies are specific to the used inversion system (or similar systems). The results may not be valid for other systems, especially those accounting for large scale or temporal biases. This needs to be discussed in the paper and also clearly stated in the abstract.

See above responses to general comments. Our results highlight the sensitivity to (sub-) regional observation bias within and outside Europe. We also show that it is difficult to develop a robust bias correction scheme before the systematic errors with both the observations and models have been properly characterized. Instead, using different observations under different assumptions, we are able to infer quite different European uptakes, which cannot be invalidated by currently available observations. So it is reasonable for us to question whether the elevated European uptake inferred from GOSAT data is real or an artefact due to small regional or sub-regional biases.

(2) Abstract: Page 1991, line 9 (P1991 L9): "We show ..." the performed studies are not suitable to show this (see general comments).

Please see our response above.

(3) Abstract: P1991 L20: "We find that ..." the performed studies are not suitable to show this (see general comments).

Please see our response above.

(4) P1993 L8: "Consequently, there is an ongoing debate ..." :
i) Here Feng et al. imply that the presented study can make any conclusions whether the results of Reuter et al. (2014) are realistic or not. This is not the case because in contrast to the used inversion system, Reuter et al. (2014) set up an inversion system that is by design in sensitive to seasonal biases, large scale regional biases, and is less sensitive to long range transport issues. As mentioned earlier, the presented study only has the potential to conclude about the used or similar inversion systems.

Our paper suggests a different explanation for the elevated European uptake inferred from the GOSAT XCO2 data, which can be not ruled out by the current observation network. Please see our response above about the impact of regional bias on our global and quasi-regional inversions.

ii) Please note that the results of Reuter et al. (2014) are qualitatively consistent with those of, e.g., Basu et al. (2013) and Chevallier et al. (2014). These references should also be cited in this context because the consistency among different inversion systems, bias correction schemes, satellite instruments, etc. shows how robust this signal is.

Our GOSAT-only inversions (which earlier results are included in Chevallier et al (2014)) also show a similar enhanced uptake over Europe. Here, we just put forward an explanation of this result.

(5) HIPPO comparison (P1995 L8-11): The performed HIPPO comparison is way too superficial to conclude which fluxes are more realistic. Please discuss/consider the following:

i) Which HIPPO campaigns have been used?

We use HIPPO-3 from 2010. We have clarified this in the revised manuscript.

ii) HIPPO measurements are mostly performed above the Pacific Ocean which requires accurate long range transport modelling to be correct when interpreting these data via inverse modelling or performing comparisons with model simulations.

We agree that (in-)validating flux inversions need to consider transport errors, representation errors, etc. Here, our comparisons are used to show that GOSAT-only inversions generate higher a posteriori model concentrations over lower latitudes than the in-situ only inversions. These higher concentrations are not supported by HIPPO observations. We have stressed this in Section 2 as well as in our concluding remarks.

iii) The signals here are probably not very sensitive to European fluxes.

The comparison is not about the observation signals to constrain European fluxes directly. Together with the comparison with CONTRAIL data over Europe, it just shows that in the air mass inflow to Europe, CO_2 concentrations from GOSAT-only inversions are higher than that from in situ inversions, and they are also higher than independent HIPPO-3 observations at middle and low latitudes. When CO_2 in air inflow is overestimated, the mass balance will dictate an elevated European uptake, even when GOSAT XCO2 retrievals within Europe are not biased themselves. This point is made more clearly by additional quasi-regional inversions presented in Appendix A.

iv) A bias comparison does not say anything. One should concentrate on spatial and/or temporal patterns.

We have added a new Figure (Figure 2) to the revised manuscript, which shows that GOSAT-only inversions overestimate the concentrations over low latitudes.

v) The HIPPO comparison should be moved to the results section.

As mentioned above, the comparison with HIPPO is not included as an evaluation of our inversions but as a background for the following study focused on European region.

P1995 L25ff: CONTRAIL comparison:
i) I can see no reason to limit the comparison to two European airports only. Please use also the other European airports.

In 2010, these are only descending and ascending CONTRAIL data available for airports in Europe.

ii) The differences shown in Fig.2 seem to be small. Please add panels showing the difference and quantify the differences (e.g., by calculating the root mean square difference) and discuss whether the differences are significant.

The purpose of this Figure is to show that CO2 mole fraction data corresponding to fluxes inferred from GOSAT data are higher than those inferred from in situ data at the beginning of 2010. We are happy to show this comparison as differences instead of absolute amounts. But considering that the agreement with aircraft data is affected by many other factors such as model transport and representation errors, and unaccounted small-scale local sources etc., it does not invalidate flux estimates directly. We have stressed this point in Section 3.

iii) As long as the manuscript aims at conclusions which may have implications also for previous inversion studies of others, the validation results of the corresponding studies should also be discussed in this context. As an example, Reuter et al. (2014) finds improved agreement with independent measurements when the satellite measurements are inverted.

Recent inter-comparisons of 8 different inversion systems by Houweling et al. (accepted for publication in *J. Geophys. Res*.), which also include our earlier results, show that a posteriori model concentrations inferred from GOSAT X_{CO2} data generally have worse agreement with independent aircraft observations than those inferred from in situ data. Some inversions based on GOSAT data have shown better agreement with TCCON network, partially because TCCON network has been used by several retrieval groups to correct GOSAT X_{CO2} biases. In the present study, we have already included TCCON data into the reference

inversion to improve the agreement with TCCON network. However as mentioned by the reviewer, considering unaccounted model errors and limited observation coverage, worse agreement with currently available observations does not necessarily falsify GOSAT data in favour of the in situ data. Our concluding remarks stress this point.

(7) Seasonal cycle (Fig.1, P1995 L13-24):
i) GOSAT has a poor coverage during winter, despite this, there is a large deviation between a priori and GOSAT a posteriori in winter. Why?

First, GOSAT has a few observations over the Southwest part of Europe. More importantly, different from Reuter et al., we have used a 5-month assimilation window, so that observation made in late Spring and early Summer were also used to constraint flux estimates during winter months.

ii) The results of Reuter et al. (2014) suggest that the largest uptake-increment occurs during the growing season when GOSAT observation conditions are advantageous and a priori uncertainties are largest. This point should be discussed.

As discussed in the manuscript, our inversion using in situ data already has a large summertime uptake due to the use of TCCON data. The size of this summertime uptake peak agrees well with previously reported GOSAT inversions from our own and other groups (including Reuter's peak values).

(8) Inconsistency with other studies (e.g., P1995 L13ff):
i) INV_ACOS and INV_UOL are GOSAT only inversions. The combined inversions result in a European sink of 0.61 GtC/year (ACOS_INS) or 0.66 GtC/year (UOL_INS). This seems to be not consistent with the results of Chevallier et al. (2014); Basu et al. (2013); Reuter et al. (2014) finding an annual sink of roughly 1 GtC/year for combined or step wise in situ plus satellite inversion.

When combining data from different sources, the flux estimates obviously depend on the relative weights given to the in situ and GOSAT data. This point has been included in in Section 4. In fact, smaller European uptake estimates have been found in several experiments when assimilating GOSAT XCO2 and in situ observations together (see Houweling et al., 2015). Also, we have included TCCON observations as additional in situ observations, and introduced bias corrections to GOSAT XCO2 retrievals. As a result, we anticipate a strong influence by in situ observations where they are available. Reuter et al. used a quite different approach.

to other inversions assimilating the same/similar in situ/GOSAT datasets (e.g., Basu et al., 2013; Chevallier et al., 2014; Peters et al., 2007; Peylin et al., 2013). Why are these error estimates so much smaller than those of, e.g., CarbonTracker which bases also on EnKF? Same for most error estimates given in Tab.1.

We thank the reviewer for spotting this mistake. We mistakenly replaced the unit convert factor of 144 (=12x12) with 199 the number of geographical regions. In the revised calculations, we also increased the a priori uncertainty by 20%. The resulting a posteriori uncertainty for ACOS inversion is now 0.19 GtC/a, which is close to 0.23 Gt/Yr by Chevallier et al (2014) who has used a variational data assimilation system. More importantly, in the revision, we have stated:

'However considering the limited spatial resolution (only 12 sub regions for the whole TransCom European region), and unquantified model transport and representation errors, we anticipate that the complete a posteriori uncertainty is larger than the value estimated by the inversion system itself , as suggested by large inter-model variations found for in situ inversions (e.g., Peylin et al., 2013)'

(9) Appendix A:

i) Omit this section or discuss and interpret the results. Just mentioning that these inversions have been performed is not enough (P1997 L9-14).

We decided that the core message should be contained in the main text while the important details about

quasi-regional inversions should be kept in the Appendix to improve the overall readability of the main text. We have also rewritten Appendix A so that it is more focused on the effect of the boundary conditions as well as the on-line bias correction with a large a priori uncertainty.

ii) Scaling the a prior errors by 2/3 enhances the influence of the a priori so that you cannot compare the results with those from the other inversion set ups. Selecting the same a priori errors will result in larger European uptake. Additionally, one can expect that the uptake will still be lower than that inferred from the global GOSAT inversions because the data gets less weight due to fewer data points (expressed by the larger a posteriori errors). I expect that the resulting error bars will overlap with those from the global inversions.

We apply the factor of 2/3 to reflect the a priori is taken from the a posteriori fluxes from INV_TCCON. We have stressed this point in the revision. Also, using an alternative factor of 1, for example, increasing the uptake by 0.08 GtC/a, which does not change our discussions significantly.

(10) Observations outside Europe (P1991 L20ff, P1997 L3ff, P1998 L16): The conclusions drawn from the MOD_NOEU and MOD_ONLYEU analyses are unclear oreven wrong. The statement that 50% or 80% of the signal comes from measure-ments outside Europe does not help to conclude if the observed GOSAT signal is real or an artefact. Due to atmospheric transport, it is clear that measurements outside Europe include information about the European carbon sink. Let us assume for a mo-ment that GOSAT had no errors and many more measurements (so that the results are dominated by the measurements and not the a priori). The concentrations of the INV_TCCON inversion are consistent with a 0.6 GtC/year sink. This is reflected by the results of the INV_ACOS_MOD_ALL. GOSAT measurements are consistent with a 1.2 GtC/year sink. If you mix real GOSAT measurements with concentrations of the INV_TCCON it is clear that you will end up with an European flux somewhere between inner/outer European measurements to the European flux. If GOSAT inversions where unrealistic due to retrieval biases, I would expect a rather little probability of having inner European bias patterns being consistent with outer European bias patterns. The fact that the fluxes of the ACOS_MOD_NOEU and ACOS_MOD_ONLYEU are basi-cally equal indicates that the inner European GOSAT information is consistent with the outer European information. This could be interpreted as indication for the GOSAT signal being real and not an artefact. This should be discussed. In this respect, see also my comments related to Appendix A (see comment (9)).

We respectively disagree with the reviewer. The monthly fluxes inferred in ACOS_MOD_NOEU and ACOS_MOD_ONLYEU are actually quite different from each other: ACOS_MOD_NOEU shows much larger uptake at the beginning of 2010, while ACOS_MOD_EU has a larger uptake during the summer. The resulting annual net uptake is just a coincidence, which has not been shown in results for another (UOL) GOSAT data set.

(11) P1998 L9-10: "...coarse coverage of in situ observations is unlikely..." I don't see how your study supports this hypothesis.

We mean that even including the GOSAT observations over eastern Europe, the joint data inversion is not far from the in situ inversion, particularly after the on-line bias correction. But we agree that the contribution from observations over East Europe need further quantification in the future. So we have removed this sentence.

(12) Monthly biases: As discussed at P1997 L14-22, monthly biases can influence the derived annual fluxes.

i) For this reason Reuter et al. (2014) simultaneously fits monthly biases. They still find a European sink of 1.0 GtC/year. Their results should be discussed in this context

First, the ultimate aim of this paper *is not* to directly invalidate results described by Reuter et al.

ii) It should be discussed that the derived potential impact due to monthly biases is only valid for the used inversion system. Other systems may be insensitive to such biases (e.g., Reuter et al., 2014).

As discussed in above responses, this study provides another hypothesis on why global flux inversions infer much larger European uptake from GOSAT X_{CO2} retrievals, and such hypothesis cannot be ruled out by

current observation network. But we agree that different systems may have different sensitivity to certain biases we presented here. So in the concluding remark, we added:
'Complicated interaction between observation and the assimilation system means that our present study does not exclude other possible causes for the elevated European uptake reported by previous researches from assimilation of GOSAT data. Instead, it highlights the adverse effects of uncharacterized regional biases in current GOSAT X_{CO2} retrievals that can attract erroneous interpretation of resulting regional flux estimates…'

iii) The conclusions (P1999L17ff) imply that the GOSAT data indeed suffer from a seasonal bias explaining 0.3 GtC/year. However, the performed analyses can only conclude on the sensitivity of the used inversion system. Even though the UoL and ACOS fluxes are similar, Fig.4 does not suggest that there is a (spatial or temporal) common bias patter in the satellite retrievals.

That is true for the West European region. However, there appears to be a generally positive bias over the tropics (Figure 2), which, unfortunately, cannot be directly confirmed by using current TCCON network. Interestingly, when we remove the impact of GOSAT data outside the European regions, the inversion based on UoL v4 data has a net European uptake very close to that inferred from the in situ data. This is one of the results that motivated us to write this manuscript.

(13) Regional inversions: P2000L13ff "Our study suggests that...": The performed analyses does not support this hypothesis. The analysis ACOS_EU and UOL_EU have an posteriori error of ±0.16 GtC/year which would be a huge uncertainty reduction over the current IPCC estimate of about ±0.4 GtC/year (Peylin et al., 2013). Additionally, the statement ignores potential advantages of regional inversions, e.g., being insensitive to retrieval biases outside Europe.

We agree that the a posteriori uncertainty from inversions tend to be underestimated by the inversion system itself, which is stressed in the revision (Section 2). We also agree that regional inversion has some advantages over a global inversion. But for regional inversions, accurate boundary conditions are required, and bias correction without proper bias characterization is also challenging (see Appendix A).

(14) P1991 L4: "Recent work has shown...": This is too general. Peylin et al. (2013) shows that in situ based inversions span over a large range with values up to about 1.5 GtC/year.

It has been changed to: 'Several recent researches …'

(15) P1991 L6: Add "and the Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY (SCIAMACHY)".

It has been added.

(16) GOSAT data: The manuscript should explain why outdated data sets have been used (ACOS v3.3 and UoL FP v4.0). The most recent versions are ACOS v3.5 and UoL FP v5.1. As an example, Reuter et al. (2014) used already ACOS v3.4. (discussion paper submitted on July 23th 2014)

This is a good point. A lot of development work has been completed on the retrievals since the publication of Chevallier et al., but we wanted to continue the scientific discussion that was motivated by this paper.

(17) P1991 L9: What do you mean by "mis-fitting"?

We have rewritten the sentence to avoid confusion.

(18) P1991 L26: Add something like "for the used (or similar) inversion systems".

We have added 'for our global flux inversions …'

(19) P1992 L13: "spatial scales < O" what do you mean?

Which means at a spatial scale of several thousand kilometres. The O denotes the "order of magnitude" and is standard mathematical usage.

(20) P1992 L23: "Surface flux estimation algorithms are particularly sensitive to ..." This is too general because the sensitivity strongly depends on the used inversion technique and the spatial and temporal correlation length of the bias patterns. See, e.g., Bergamaschi et al. (2007); Basu et al. (2013); Reuter et al. (2014) for inversion techniques being less sensitive to specific bias patterns and Miller et al. (2007) for a discussion of spatial error correlation lengths.

We have changed the sentence to: 'many inversion systems are …'. But, as discussed in Appendix A, developing a robust bias correction system is non-trivial, and required proper characterization of bias errors.

(21) P1994 L19: "Including TCCON..." How large is the additional error reduction due to TCCON? Please add error estimates for the given fluxes (0.47 and 0.56 GtC/year). If it is low, the "true" European sink could be much larger than 0.56 GtC/year (assuming

We have changed the manuscript accordingly. For Europe, inclusion of TCCON sites reduced the uncertainty by about 15%.

(22) INV TCCON: The "reference" inversion set-up uses in situ and TCCON measurements. Why not using an in situ only inversion as reference?

We chose to include TCCON data because they improve the estimate of summertime peak uptake.

i) In this case, you could use TCCON as additional independent validation data set.

We want to define a reference system constrained by in situ data that was as accurate as we could make before comparing it against GOSAT data.

ii) TCCON may have station-to-station and/or seasonal biases of about 0.4 ppm Wunch et al. (2011).

This is a good point, although smaller bias is expected for GGG2014. In the revision, we have taken into account these errors, by adding 0.5 ppm to TCCON observation errors (See Section 2).

i) "While the GOSAT inversions suffer from ...": Please add "larger" because in situ measurements also may have observation errors.

Good suggestion, and 'larger' is added.

What do you mean by aggregation errors. In the sense of Kaminski et al. (2001); Engelen et al. (2002), aggregation errors are less an issue for GOSAT because of the denser spatial sampling and the fact that seasonally no "hard constraints" are used.

The 'aggregation errors' is about the effects of the seasonable observation coverage. It is changed to: '…issues from the seasonal coverage of higher latitudes'

(24) P1993 L26: Within the abstract you mentioned to use an EnKF now its an ETKF?

ETKF is the numerical algorithm to solve the optimal fluxes and the associated uncertainties. To avoid confusion, we replace it with EnKF.

(25) P1994 L13: Please add information on the used a priori error correlations.

We have revised the manuscript accordingly in Section 2 to provide detailed information about assumed

temporal and spatial correlation of the a priori uncertainty.

(26) P1997 L27: What do you mean by "control variables"? Have you added a monthly TRANSCOM wide bias to the state vector? If so have you included some month-to-month a priori error correlations?
We make it more clearly in the revision:
'We also include monthly GOSAT X_{CO2} regional biases over 11 TransCom land regions [Gurney et al., 2002] as parameters to be inferred together with surface fluxes from the joint assimilation of in-situ and satellite observations …'
No correlation is assumed, because we don't know.

(27) Table 1: Add a priori flux and uncertainty.

Good suggestion. We add in the caption of Table 1:
'Except INV_ACOS_INS_DBL_ERR and INV_ACOS_DBL_ERR, the aggregated European annual uptake of the a priori fluxes is -0.1±0.52 GtC/a.'

(28) Bibliography: Please check all citations. As an example, Deng et al., 2013, is Printer-friendly Version

We have updated citations.