Review Report one

Major comments:

'The authors have improved their text remarkably well, except about my first remark. They still do not comment on the divergence between their experiments and those of Reuter et al. They claim to have "put forward an alternative hypothesis and to have shown that realistic levels of bias in the GOSAT data can result in an erroneous elevated uptake over Europe" (abstract), but Reuter et al., including half of the current team, claimed that their system was "insensitive to large-scale retrieval biases " (abstract of Reuter et al.) and was "solely dependent on regional (medium-scale) gradients" (introduction of Reuter et al.). They also investigated the effect of possible regional biases and concluded that "retrieval biases in mean wind direction are unlikely to explain the observed carbon sink" (Appendix B of Reuter et al.). In other words, rightly or not, Reuter et al. already ruled out the hypotheses brought forward by Feng et al., that are therefore not as new as claimed by Feng et al.

Such experimental disagreements happen, but they have to be clearly stated and commented. The authors need to explain why the reader should trust the current results at least as much as those of Reuter et al., and what can be learned from this divergence between two sets of results.'

We'd like to thank the reviewer again for the helpful comments. Reuter et al have used a regional flux inversion approach, while this manuscript is mainly based on a global flux inversion system. In the Appendix A, we have discussed issues associated with the quasi-regional flux inversions, before and after introducing a simple on-line bias correction scheme with a large prior uncertainty. However, many prospects of our quasi-regional flux inversions are different from one used by Reuter et al., and we cannot simply conclude on their approach. Instead our results suggest that without proper characterization of observation biases, it is non-trivial to develop a robust on-line bias correction scheme without the significant side effects, such as unintended information loss, and enhancing sensitivity to undetected biases (Appendix A). So in the revision (abstract and conclusion), we have added sentences to highlight the need to develop more robust inversion systems.

.

Minor:

l. 22: this this -> this
Correction is made

l. 243: there a few -> there are few
Correction is made.

l. 248: observations -> observations alone

Modification is made.

Review Report 2

1. Overall impression

'...Despite the improvements from the initial version, I find the title and abstract of the present manuscript problematic. The title asks a provocative question, and the abstract suggests that the reader will be rewarded with an answer (yes, it is an artefact, 60-90% of which is coming from possible biases in XCO2 outside Europe), or at least a compelling argument, one way or the other. The body of the paper, however, does not answer the question asked in the title. From my reading, the overall message of the paper seems to be that small biases in XCO2 can result in large errors in estimated surface fluxes, which is not a new or surprising message in this field any more. The paper presents several experiments to demonstrate this specifically for European fluxes, which have not been published before. Therefore, a more accurate title of the paper would have been "Sensitivity tests to probe the enhanced European sink in GOSAT XCO2 inversions", or something along those lines. I like those sensitivity tests, and I think they should be in the published literature for reference '

First, we would like to thank the reviewer for providing their detailed review and recommendations.

'However, they should not appear under the current title and abstract, which is misleading. My concern about the title and abstract is not merely pedantic. I have discussed this paper (much before it came to me to be reviewed) with multiple carbon cycle colleagues who are aware of the debate surrounding the European sink, but are not atmospheric inverse modelers themselves. Their first reaction has always been something akin to "I see, this group has figured out that satellite data are to blame... '

The main aims of this manuscript are to describe another possible explanation on the elevated European uptake inferred from GOSAT XCO2 retrievals, and to emphasize that more research is required to conclude robustly whether previous reports of elevated uptake of CO2 over Europe describe a real phenomenon. The original title was just used to highlight that currently we do not know for certain whether the elevated uptake is real or an artefact.

In an attempt to explore this question we have investigated the sensitivity of European CO2 flux estimate to possible bias in current GOSAT XCO2 retrievals both inside and outside Europe. We find that for a global flux inversion system, a large portion of the elevated uptake over Europe can be related to the assimilation of GOSAT XCO2 retrievals outside the European region, with a pattern similar to the effect from over-estimated CO2 influxes to Europe. A regional flux inversion, however it is posed, will always rely on prescribed boundary conditions. We also find that GOSAT XCO2 data within Europe have a relatively small contribution to the elevated European uptake, which is explained by small, time-varying, sub-regional biases.

But following the reviewer's suggestion, we now use a new title.

'...But then why do they say in the abstract that they can explain only 0.18 GtC/a from XCO2 biases, when the enhancement they need to explain is 0.62 GtC/a? '

By using multiple inversion experiments (see Table 1 and the more detailed reply to a related question below), we have demonstrated that a large portion of the elevated uptake of CO2 can be related to the assimilation of GOSAT XCO2 retrievals outside Europe, with a pattern similar to the effect of overestimated CO2 inflows into Europe. Our results (as do others, e.g. Houweling et al., 2015) also show that assimilating GOSAT XCO2 retrievals usually results in a poorer agreement between posterior model concentration and the independent HIPPO measurements, particularly over lower latitudes where models typically have a large positive bias.

On the other hand, when GOSAT XCO2 retrievals outside EU are replaced with the model simulations (INV_ACOS_MOD_NOEU and INV_UOL_MOD_NOEU), or are not assimilated (INV_BD_TCCON in Appendix A), the inferred European uptake of CO2 becomes much smaller. The remaining uptake enhancement can also be explained by small, time-dependent sub-regional biases in XCO2 retrievals over Europe.

'...The numbers are different in the revised manuscript (which is a point I'll get to later), but the same concern about the wording of the title and abstract remains. By my reading, the content of the paper is not about disproving Reuter et al. (2014), nor is it about proving that the enhancement in the European sink is a result of biases in GOSAT XCO2. It is about demonstrating that certain spatio-temporal patterns of XCO2 biases – realistic or otherwise – could lead to an overestimate of the European sink in an inversion. The title and abstract should be changed to reflect this. The revised title may be less provocative, but it will also be more accurate …'

The different numbers were caused by using an enlarged prior uncertainty in the last revision, as suggested by the reviewers of the previous draft.

We agree with the reviewer that we do not have sufficient observations to directly support the existence of these biases, but neither are there are sufficient measurements to prove elevated CO2 fluxes over Europe are a real phenomenon. Exploring alternative hypotheses is part of the scientific research. We believe that our analysis is consistent with information we have, highlighting the importance of further studies on the performance of GOSAT XCO2 retrievals as well as on the performance of inversion systems both within and outside the European region for reaching a consistent estimate of the European uptake of CO2.

2. Major point-by-point comments

'...Line 30 The annual uptake for INV_ACOS and INV_UOL were 1.20 GtC/a and 1.16 GtC/a respectively in the first version of the paper. What changed to make them both 1.4 GtC/a in the second version? It's not obvious that the authors changed anything in their inversion setup, but the uptake changed by 31% of the enhancement they were trying to explain in the first version. That's troubling...'

The revised manuscript has addressed suggestions from the previous reviewers. For example, in the revised manuscript we recalculated the posterior fluxes using a substantially increased prior flux uncertainty (see section 2 and our replies to comments in the open discussion). As a result, the uptake inferred by GOSAT-only inversions has increased significantly (by about 0.2 GtC/a), as pointed out by this reviewer. Such sensitivity is indeed part of the motivation for us to question the causes behind the elevated uptake of CO2 over Europe.

Line 199 INV_ACOS gets a European sink stronger by 0.82 GtC/a than INV_TCCON. The authors substitute ACOS XCO2 outside Europe with the posterior CO2 field from INV_TCCON, and this gives a European sink stronger by 0.3 GtC/a than INV_TCCON. Hence the authors conclude that 0.3 GtC/a of the 0.82 GtC/a strengthening is due to biases in ACOS XCO2 inside Europe, while the remaining 0.52 GtC/a strengthening is due to biases outside Europe. This logic behind this partitioning of the enhancement to XCO2 inside and outside

Europe is flawed, because corrections to surface fluxes are not additive in the datasets assimilated. In other words, if dataset B yields an enhancement of X GtC/a over dataset A, and dataset C yields an enhancement of Y GtC/a over dataset A, then datasets B + C do not necessarily yield an enhancement of X+Y GtC/a over A. The simplest illustration of this is to consider (a) a GOSAT-only inversion, (b) an in situ-only inversion, and (c) a joint GOSAT + in situ inversion, where A is the null data set, and any surface flux estimate is an "enhancement" over the prior. Flux estimates from (c) are not sums of estimates from (a) and (b).

As shown in Table 1 and Figure 4, contributions of the GOSAT XCO2 retrievals outside EU to the elevated uptake are indeed assessed through multiple experiments. For example, the XCO2 data sets assimilated in experiment INV_ACOS_MOD_ALL, INV_ACOS_MOD_NOEU, and INV_ACOS_MOD_ONLYEU have the same observation coverage, averaging kernels and observation errors as that for the actual GOSAT-only inversion INV_ACOS. The only difference between those 3 inversions is the assigned values of their XCO2 data set:

1) INV_ACOS_MOD_ALL. They are all from model simulation. Here we generate model GOSAT XCO2 by sampling posterior model concentrations from INV_TCCON at the locations of actual ACOS XCO2 retrievals, and converting the resulting profiles into XCO2 using the averaging kernels from ACOS retrievals;
2) INV_ACOS_MOD_NOEU. XCO2 values outside EU are taken from model simulation as INV_ACOS_MOD_ALL, and within EU we use the actual ACOS retrievals as INV_ACOS;
3) INV_ACOS_MOD_ONLYEU. Outside EU values are from model simulation, while inside EU values are from actual ACOS retrievals.

The resulting European uptakes of CO2 are:

1) 0.64 GtC/a for INV_ACOS_MOD_ALL, which is close to the reference in-situ inversion INV_TCCON of 0.58 GtC/a;
2). 0.88 GtC/a for INV_ACOS_MOD_NOEU;
3). 1.17 GtC/a for INV_ACOS_MOD_ONLYEU.

When actual XCO2 retrievals are assimilated both within and outside EU, INV_ACOS infers a net European uptake of 1.40 GtC/a. Those experiments together clearly demonstrate the effects just due to different XCO2 value assimilated within or outside EU. Large contribution from XCO2 retrievals outside EU is also found in our quasi-regional experiments in Appendix A. To address this reviewer concern we have added a sentence in section 3 to emphasize that conclusions are drawn from several experiments.

Line 215 Adding a 0.5 ppm bias in Feb-Apr (Jun-Aug) yields a reduction in uptake by 0.1 GtC/a (0.15 GtC/a). None of those reductions are close to the number the authors are trying to explain as being due to biased XCO2 inside Europe. Even if they were, it wouldn't mean that the enhanced European uptake is caused by seasonal 0.5 ppm biases. There are lots of different ways to change the observations and get the same answer in a source-sink inversion. As it stands, this experiment simply shows that inversion results are sensitive to small biases in XCO2 . Same comment applies to lines 212-218; just because adding 0.5 ppm CO2 has a certain effect does not mean that for the real data the same effect is the result of a 0.5 ppm bias.

The aim of these sensitivity experiments is to demonstrate how the inversion system responds to the assumed bias. We agree with the reviewer that they do not prove there are (or not) such biases. However if the experiments show that the inversion results are sensitive to small biases, it is natural for us to question whether they are reliable as we are unable to validate XCO2 retrievals to a high accuracy over most geographical regions by using the current validation network.

' Line 225 The way the authors have set up the biases in XCO2 to be optimized is unphysical. Any possible bias in GOSAT XCO2 is highly unlikely to vary with TRANSCOM regions. Rather, it may vary with geo-physical parameters on which the retrieval code depends, such as solar zenith angle, surface albedo and type, topography, aerosol loading, etc. A parameterization of the bias where it depended on those variables would have been much more physical and believable (e.g., the overall land-ocean offset of Basu et al. (2013)). Instead, the way the authors have set up INV_{ACOS,UOL}_INS, the bias parameters vary at the same spatiotemporal scales as the fluxes they are trying to estimate (monthly, over TRANSCOM regions), and there is nothing in the atmospheric CO2 data that can distinguish between the two.North America or Europe is ∼0.5 ppm, and even by the authors' own accounting, a 0.5 ppm bias can have very significant impacts on surface fluxes (this is the point of several of their experiments). So what the authors are doing by estimating so many bias parameters – each of which can deviate up to 1.5 ppm from zero – is essentially throwing away any regionally coherent information in GOSAT XCO2 whenever it doesn't agree with the posterior field of INV_TCCON. For this reason, I'm not at all surprised that their estimates from these two inversions (0.62 GtC/a and 0.67 GtC/a) are so close to INV_TCCON; they've simply bias corrected away most features in GOSAT XCO2 which did not agree with INV_TCCON. Why even bother assimilating GOSAT XCO2 , in that case?

As section 5 explains, our online approach is mainly used as consistent way to derive a simple estimate of systematic differences between the model and GOSAT XCO2 at regional or sub-regional scale. We focus on those systematic differences over Eastern or Western Europe, which are shown to be small and time-dependent but be able to explain most (up to 0.2 GtC/a) of the remaining uptake enhancement. This highlights the challenge we face to rule out elevated uptake estimate caused by small biases.

In our approach we have 199×12 flux variables compared to a total of 12×12 variables for regional XCO2 biases. In section 4 and Appendix A we have cautioned the reader about using the on-line bias correction, which may lead to the loss of real information particularly when prior uncertainty for XCO2 bias is set to be very large. The assumed uncertainty of 0.5 ppm for monthly biases at regional or sub-regional scale is generally consistent with the deviations revealed by recent inter-comparisons between XCO2 retrievals and TCCON network or between XCO2 retrievals and model simulations (see for example Lindqvist et al., 2015; Kulawik et al., 2015; Chevallier et al., 2015).

We always treated the inferred value as an averaged difference between the inversion system and GOSAT XCO2 retrievals over different geographic regions. We have not stated that the estimates from INV_ACOS_INS and INV_UOL_INS are the closest one to the true surface fluxes.

We agree with the reviewer that although mean differences at a certain spatial scale are frequently used in various comparisons between model and observations, it would be much better to relate these differences to certain geophysical variables. However we are discussing the elevated uptake inferred from (typically) bias-corrected GOSAT XCO2 retrievals (e.g., Chevallier 2014; Houweling et al., 2015). It is therefore unclear to us how the remaining systematic errors are related to geophysical variables (Chevallier et al., 2015). It is also worth stating that biases are estimated by the data providers typically on the basis of an assumed dependence on geophysical variables. Using a parameterized bias correction still risks the loss of real information because many geophysical parameters (e.g., surface albedo and solar zenith angle) have seasonal cycles that are correlated with the natural biosphere.

Line 228 I do not understand the authors' choice of splitting up the bias estimate over eastern and western Europe. They have not presented a physical argument for why they think a spatially uniform bias over eastern Europe would be different – and uncorrelated – from a spatially uniform bias over western Europe. Their choice, in fact, might have done more harm than good. Reuter et al. (2014) found that the enhancement of the European sink could come from the east-west gradient in GOSAT XCO2 within Europe, even if the average GOSAT XCO2 over Europe as a whole was unbiased compared to in situ data based inversions. By separately estimating biases in eastern and western Europe in the paper, the authors are essentially guaranteeing that their INV_ACOS_INS inversion will not interpret this gradient as a signature of surface fluxes, even if such a gradient really is present in the atmospheric CO2 field.

As discussed in our responses above as well as in the manuscript, we have stressed that our simple approach is mainly used to estimate systematic differences between model and XCO2 retrievals over two sub-regions East and West Europe. We have *not* assumed the biases are uniform within each sub-region: we treat the derived values as the averaged differences. We focus on the point that small, time-dependent biases can explain most (up to 0.2 GtC/a) of the remaining uptake enhancement in order to highlight the challenges of detecting and removing observation biases.

We use our on-line bias correction with caution, and have clearly pointed out in the manuscript that our bias correction can result in loss of real information, which may include part of the real gradient mentioned by the reviewer. Appendix A presents the side effects when large a priori error for bias is assumed in our quasi-regional inversions (Figure A2 and Table A2). For our joint data assimilation system the difference with or without using our simple online corrections is around 0.1 GtC/a (for INV_UOL_INS) or 0.15 GtC/a (for INV_ACOS_INS) (see section 5). We have not assumed the East-West gradient will not affect the flux inversions, but the effect from information loss appears limited in our inversion system.

'Line 251 INV_ACOS_INS with 0.01 ppm uncertainty on the bias parameters (INV_ACOS_INS/0.01) yields a sink quite close to INV_TCCON and far away from INV_ACOS. Yet by design, posterior XCO2 from INV_ACOS_INS/0.01 should match GOSAT XCO2 everywhere, as well as in situ CO2 . This means that GOSAT XCO2 – biased or not – are consistent with a much lower (0.77 GtC/a) European sink than that found by INV_ACOS. So whatever biases may exist in GOSAT XCO are not responsible for the 0.82 GtC/a strengthening of the European sink, nor is such a large European sink necessary for explaining the observed XCO2 . Where is my reasoning wrong?

This is an important point, because this finding – a joint inversion yields a sink closer to an in situ only inversion – is unlike what previous studies found. E.g., Basu et al. (2013) found that the European sinks from their "Flasks + GOSAT" and "GOSAT" were virtually identical, and ~0.75 GtC/a stronger than their "Flasks" inversion. To resolve this question, I would like to see a plot of the difference between posterior XCO2 – at GOSAT sounding locations and convolved with GOSAT averaging kernels – from the INV_ACOS_INS/0.01 and INV_ACOS inversions. That difference is the manifestation, in XCO2 space, of 0.63 GtC/a flux from Europe. To what extent the two inversions match GOSAT XCO2 is, of course, defined by the error settings on their in situ vs GOSAT observations. The authors have already presented their errors for GOSAT XCO2 and TCCON XCO2 , so only the errors assumed for point samples is

missing from the manuscript. I would like to see the errors they assumed for near surface point samples (flask, tower, ship, ...).'

We have read carefully but failed to understand the reason behind the following comment: "*whatever biases may exist in GOSAT XCO are not responsible for the 0.82 GtC/a strengthening of the European sink, nor is such a large European sink necessary for explaining the observed XCO2.*", just because the European uptake estimate from the joint in-situ + GOSAT inversion is closer to the in-situ only inversions.

Several groups have shown that joint inversions result in significantly less uptake than the GOSAT-only inversions (Houweling et al., 2015). The joint data assimilation system will comprise or mitigate between the in-situ and GOSAT data based on their estimated uncertainties. Considering we have added the TCCON network as part of in-situ data set we expect a strong influence from the in-situ data on estimating the European uptake of CO2, particularly over the winter and early spring period when coverage by GOSAT is poor. Our results may also indicate the joint assimilation system is less sensitive to biases in XCO2 data (as they exist) due to the inclusion of the in-situ observations.

'Line 267 The authors seem to have performed an inversion (not in their Table 1) where they've corrected XCO2 over Europe by their estimated bias correction and left the other XCO2 untouched, which yielded them a reduction in the annual uptake by 0.2 GtC/a. Why didn't it yield a reduction by 0.3 GtC/a, which is what the authors earlier said was the impact of biased soundings inside Europe? This seems to be an inconsistency.'

As stated in the above replies, we have estimated a sub-region bias using a simple assumption, which was shown to be able to explain most part of the remaining extra uptake. We have never excluded other causes for the reported uptake enhancements, such as the varying observation coverage, and the model transport error etc. For example even when model XCO2 values are assimilated, INV_ACOS_MOD_ALL has a slightly higher uptake than INV_TCCON. Also it is highly possible that at smaller scales local XCO2 bias may be higher the region mean, resulting in more uptake than we estimate using a simple approach.

'Line 298 The authors seem to implicitly assume that their INV_TCCON result is the "more correct" one. It is from this viewpoint that they say that because INV_ACOS_DBL_ERR yields a sink of 1.61 GtC/a compared to 1.4 GtC/a of INV_ACOS, the problem must be biased XCO2 . I can think of another possibility. What if the CO2 field over Europe is consistent with a stronger European sink, but INV_ACOS doesn't get there because the prior is too tightly constrained, INV_TCCON doesn't get there due to sparse sampling of the surface layer, and INV_ACOS_INS doesn't get there because too many bias parameters are being optimized which are functionally identical with the signature of surface fluxes?'

The question about INV_ACOS_DBL_ERR is not only because of the larger annual net uptake, it is also because, as shown in Figure 4, that the resulting net biospheric flux at the beginning of 2010 becomes more negative (i.e., higher net uptake) from an already low value by INV_ACOS. A strong net biospheric uptake during the early months of 2010 is suspicious at best. In addition, as mentioned in above replies, we have 199×12 flux variables, compared to a total of 12×12 variables for regional XCO2 biases. Also, doubling prior flux errors means that flux estimates become easier to be adjusted by the data assimilation . In addition, as we have mentioned in the

Section 2, doubling the a priori error only slightly changes the uptake by INV_TCCON (by about 0.09 GtC/a), where no on-line bias correction is involved

'I'm not saying that the European sink is 1.61 GtC/a, but that the authors haven't convinced me that it isn't. This particular comment is representative of my general problem with this paper, which I mentioned at the beginning. Namely, the authors perform sensitivity tests which are interesting, but which do not resolve the question they purport to answer in their title and abstract.'

The title conveys the message that there is a question to be answered instead of simply assuming one hypothesis is correct. We hope that our replies are able to clarify some misunderstandings. Such discussions always help improve our manuscript, and necessary changes are made to stress the basic message from our paper that more accurate and precise observations are required to test XCO2 retrievals as well as the inversion system. This message will not change with newer space-borne sensors.

3. Minor points

' Line 21 To my knowledge, only one group out of all the groups referenced in the paper has claimed that the enhancement of the European sink could be real, namely Reuter et al. (2014). All the other publications either label it a possible artefact (Chevallier et al., 2014) or are silent (Basu et al., 2013; Deng et al., 2014; Nassar et al., 2011; Takagi et al., 2014). So instead of saying "some groups", the authors should clearly mention that publication/group. Even that group, it should be noted, stopped just short of concluding "this is a real signal", instead saying (to paraphrase) "If it's an artefact, it's not obvious from the thirteen tests we have run, so we should revisit the bottom up European sink.'

We just tried to point out that that people have different opinions on elevated European uptake reported previously. We agree with the reviewer, and rewrite the sentence as:
some recent researches suggest …

Line 72 Please replace "emission" by "source". In typical usage in literature, the former includes, e.g., fossil fuel combustion but not ecosystem respiration.

Change is made.

Line 96 What is the assimilation time window for the EnKF setup? Assimilating satellite data may need a longer time window than surface data, since XCO2 has more far field influence. Have the authors considered that?

Like our previous experiments (Chevallier et al., 2014), we have used a 5-month lag window to assimilate in-situ observation and GOSAT XCO2 retrievals (Feng et al., 2009). Recently we have made new experiments with higher spatial resolution and shorter (4-month) lag windows. The results are consistent with those presented in this manuscript.

'Line 105 Olsen and Randerson (2004) contains the formulation for distributing monthly CASA fluxes to three hour windows. Please cite the appropriate CASA reference for the actual monthly fluxes used.'

As stated in Section 2, we are using the 3-hourly CASA fluxes, instead of the monthly value

'Line 125 Calculating the error in TCCON XCO2 as the variation of XCO2 about the daytime mean seems overly pessimistic. The "error" of an observation in data assimilation represents the part of its variability that the forward model cannot capture. In the case of TCCON XCO2 , the forward model probably captures some of the variation in XCO2 between 9:00 and 15:00 LT, which should therefore not be part of the error assigned to the observation.'

We agree that forward models are able to capture part of fast temporal variations. But we have not fully quantified their fittings with the hourly TCCON data over different locations and different time period. So we have used a conservative approach.

' Line 136 Evaluating posterior concentrations against HIPPO data for this paper does not seem relevant, since HIPPO has no data over Europe. I understand that the authors want to show that INV_ACOS and INV_UOL put too much CO2 over the tropical Pacific, but they haven't made it clear why that's important in this context.'

See above replies for the related major points. Current study suggests that proper description of CO2 influxes into Europe is important for us to infer a reliable European uptake. This comparison with HIPPO is used to show that posterior model concentrations from the GOSAT only inversions (INV_ACOS and INV_UOL) do not have a consistently good agreement with independent observations outside EU, with significant overestimation of CO2 concentrations over lower latitudes.

Line 145 I do not see how INV_ACOS and INV_UOL fitting GOSAT XCO2 better than INV_TCCON is any indicator of performance. The former two, after all, ingest GOSAT XCO2 , so they should fit GOSAT XCO2 better! At best, this is proof that the machinery of the inversion system works.

It is important to show that the inversion system fit the assimilated observations properly.

Line 175 Small local sources/sinks (please change "emissions" to "sources") are extremely unlikely to show up in CONTRAIL and HIPPO comparisons, since those measurements – being at high altitudes or remote locations (or both) – have diffused and large surface footprints.

Here the HIPPO and the CONTRAIL taking off/landing data are taken from a vertical range from hundred meters to several thousand meters above surface. They can be affected by local sources.

'Line 188 The prevalent flow pattern in the free troposphere over Europe in May (indeed,

through all of spring) is from west to east, as can be seen in figure C1 from NCEP/NCAR reanalysis. If, as the authors say, GOSAT XCO2 soundings outside Europe result in an unrealistically high inflow of CO2 in the hybrid run, then how can it result in an east-west gradient over Europe that is positive (0.16 ppm), as the authors claim? If anything, the east-west gradient should be negative, because the only way the excess CO2 could be flowing in – consistent with prevalent wind patterns – is from the west. The additional sink over Europe in May (all inversions in Figure 1 in the paper) should only make this negative gradient stronger.

We have checked and confirmed our results. It is an interesting question, although we do not see a simple relation between our result and the 30-year mean wind fields at 600-700 hPa. Vast regions over west and East Europe can be affected by air influxes of different origins (and hence with different concentrations). Air mass at different altitudes can also have different (ages and) origins as well. In addition, as shown in Figure 2, we have compared the hybrid simulation and the one forced by posterior fluxes from INV_TCCON over the Amsterdam and Moscow airports (please note that difference scales were used in its bottom and top panels). In April, the hybrid simulation over Amsterdam airport is about 0.79 ppm higher than the result for INV_TCCON, while the increase by the hybrid simulation for Moscow Airport (at the 'Eastern Europe') is much larger (1.11 ppm). On the other hand, in May, the increase is about 1.50 ppm for both the two airports. It would be interesting for future study to track down the origins of the air mass at different altitudes over different horizontal locations. But for our current study, the major point is that the internal gradient can be affected by different lateral boundary conditions.

'Line 194 The fact that INV_ACOS_MOD_ALL reproduces INV_TCCON reflects that the machinery of the inversion system ingesting XCO2 data works. Not sure how it reflects the seasonal variation in GOSAT coverage.'

See the response to a related major comment. INV_ACOS_MOD_ALL is part of the experiment set to evaluate impacts from GOSAT XCO2 data within and outside Europe. The model simulations are sampled at the locations of the actual GOSAT XCO2 retrievals, which reflects seasonable variations of the GOSAT coverage.

Line 199 Why did the authors not present the converse experiment of INV_ACOS_MOD_NOEU, namely, INV_TCCON posterior XCO2 inside Europe, actual GOSAT XCO2 outside Europe?

Such experiment has been done and clearly included in Table 1.

Line 225 In the setup with bias parameter optimization, does a positive bias mean that GOSAT XCO2 is too low, or too high? As in, what is supposed to be zero if all the data are fit perfectly, XCO2 model –XCO2 GOSAT − bias, or XCO2 model – XCO2 GOSAT + bias? Please write down the explicit equation(s) relating the bias-corrected and uncorrected XCO2 .

As we are discussing bias in current XCO2 retrievals, positive value means GOSAT XCO2 data is higher than the true value. The real problem is that we don't have sufficient observations to detect such biases with high accuracy.

Line 231 "The main advantage of the on-line bias estimation is that the uncertainties associated with errors in flux estimates can be partially taken into account." It is not clear to me what the authors are saying here. Flux uncertainties taken into account for what? Posterior or prior flux uncertainties?

Simply speaking, this approach takes into account the posterior flux uncertainty (after assimilation of in-situ observation and the 'bias-corrected' GOSAT XCO2 data), when comparing model and GOSAT XCO2 data.

Line 235 Representation errors are errors made because the model cannot represent concentrations at sub-grid scales, so they're relevant at small scales, not over continental scales like the authors say.

We generally agree with the reviewers' opinion. But coverage over Europe is very coarse during the winter and early spring, and the representation errors can still affect the inter-comparisons.

Line 275 Same comment as before, i.e., the outcome of INV_ACOS_OUT_0.5 ppm does not imply that soundings outside Europe are high biased, just that flux estimates are sensitive to small biases in XCO2 .

See our response for the major comments.

Review Report 3:

This paper is clearly intended to challenge the conclusions of the regional study performed by Reuter et al. (2014). They do so by arguing that the large European uptake is primarily due to measurements outside the European domain, which couldn't be assessed properly by a regional modelling framework (as that of Reuter et al.). They then "explain" the rest of the discrepancy by deriving simple bias scenarios that bring the fluxes from the GOSAT inversions into agreement with those from the surface-based inversions. They back away from suggesting that these biases are actually there, however, and only suggest that there are insufficient measurements to prove or disprove any of these results, which tells us very little indeed.

In this manuscript, we report how observation biases of data collected within and outside Europe can affect the estimate of European CO2 uptake. Sensitivity test is an important part for fully validating top-down flux inversions. It helps understand the different results when different data sets are assimilated. Based on a set of inversion experiments (Table 1), we suggest a simple explanation of the elevated uptake of CO2 over Europe. Our explanation is generally consistent with the information we currently have, highlighting the challenge we face when interpreting real data.

Insufficient observation coverage is a general challenge for the community to assess GOSAT XCO2 retrievals as well as to assess model simulations (Lindqvist et al., 2015; Chevallier et al., 2015; Houwelling et al., 2015) properly. Our study itself does not address the gap in observations, but it does highlight the need for validating space-borne observations of XCO2 under more and different observation conditions and at different temporal and spatial scales, and the need for developing more robust inversion systems.

From a logical point of view I found the two arguments to be rather at odds with one another: if so much of the excessive uptake is due to measurements outside of Europe, wouldn't one suspect a bias in the measurements outside of Europe, rather than in the region with the densest coverage for TCCON measurements? The HIPPO analysis in Figure 2 shows generally good agreement with all simulations around the European latitudes, but large discrepancies in the tropics. A more interesting and evidence-backed question might be how performing a bias correction in the tropics (to match HIPPO, or based on the newer TCCON sites) might impact the European fluxes

Bias of XCO2 data outside EU does not exclude bias of XCO2 data within EU**.** Also, the shown agreement with HIPPO data also does not exclude small XCO2 biases over EU. We agree that it is critical to address possible biases at lower latitudes, which indeed forms part of our argument. However, there are insufficient independent observations that will allow us to achieve the required validation or bias correction in a robust way.

Our joint data assimilation experiments (INV_ACOS_INS and INV_UOL_INS) represents an effort to detect and remove bias in XCO2 retrievals over different regions, including tropical lands, with the help of the model and in-situ observations. However coverage of in-situ observations over the tropics is poor.

As suggested by a previous reviewer (Dr. David Baker), another observation set available over tropical regions is the glint observations over oceans by the same GOSAT instrument. Our preliminary results show that for 2010 the posterior model concentrations from INV_ACOS are 0.47 ppm higher than the ACOS glint retrievals of version 3.4 over 0-30^{0}N, with a standard deviation of 1.26 ppm. For comparison, the mean deviations for INV_ACOS_INS (INV_TCCON) is 0.1 ppm (-0.37 ppm), with a lightly larger standard deviation of 1.30 ppm (1.36ppm). Although these comparisons support our assessment presented in this manuscript, we acknowledge that GOSAT glint data itself are neither properly validated nor fully independent.

'It is not news that small systematic errors in total column measurements can lead to large biases in retrieved fluxes. A number of studies have already tested this, including Houweling et al. (2010), which was notably missing as a reference. However unless there is some physical reason to suspect that the bias might actually have the structure in time and space that the on-line bias correction produces, it seems more like an assessment of residuals in atmospheric mixing ratios between the two runs rather than solving for a physically meaningful bias. I was also curious as to how realistic the pattern of bias corrections looked: is it very noisy, or rather continuous? Is it correlated with anything that one would expect to cause a retrieval error? Figure 5 shows that there are significantly larger bias corrections during the growing seasons over Eastern Europe than Western Europe, although the former is certainly less well-constrained by measurements. Without this sort of analysis of how plausible such an error is, it seems more like a numerical fact rather than a hypothesis: suppose z is a function of x and y; if we add so much to x and take away so much from y, we can make z approximately equal to w. This doesn't, however, tell us anything about the underlying processes. One could just as easily define the error in PBL height that is required to bring the surface-based fluxes into agreement with GOSAT-based fluxes: it doesn't mean that those errors are real.'

The impacts of biases on CO2 flux estimates have been studied previously. Our paper is focused on how biases in current GOSAT XCO2 retrievals would affect European uptake estimate. As section 5 explains, our online approach is mainly used as consistent way to derive a simple estimate of systematic differences between the model and GOSAT XCO2 at regional or sub-regional scale. It is based on a simple assumption that these systematic deviations are time and region-dependent, which are supported by recent inter-comparison of GOSAT retrievals with TCCON or with model simulations (see for example Lindqvist et al., 2015; Kulawik et al., 2015; Chevallier et al., 2015).

On the other hand, this manuscript discusses the elevated uptake inferred from (typically) bias-corrected GOSAT XCO2 retrievals (e.g., Chevallier 2014; Houwelling et al., 2015). It is therefore unclear to us how the remaining systematic errors are related to geophysical variables.

Compared to GOSAT only inversions, the joint inversions generally have a better agreement with independent HIPPO data, (as well as a better agreement with the glint data of ACOS v3.4 retrievals as mentioned in above response). But we don't consider them as the best flux estimates. Instead we mainly focus on systematic differences over Eastern and over Western Europe that are shown to be small and time-dependent but can explain most of the additional European uptake.

'Given the transport errors that clearly trouble all inversion systems (see the poor agreement of all inversions in Figure 2 with HIPPO over high austral latitudes as one example), on-line bias corrections seem questionable at best, especially when they are not constrained by any physical drivers. In contrast to this, the retrieval teams have put much time and effort into deriving bias corrections based on independent (TCCON) measurements. Unless you can demonstrate that the on-line bias corrections produces better agreement with independent measurements, it seems better left alone. I agree with other reviewers that it would have been better to leave

TCCON out of the reference inversion, and use it for validation instead (although it is, strictly speaking, not entirely independent from the bias corrections of the retrieval teams). Ignoring this suggestion was a mistake, in my opinion.'

In the above replies we stress that the on-line bias correction is mainly used to consistently estimate the systematic differences. The important information for us is that systematic differences over Eastern and over Western Europe are generally small and time-dependent, but they can explain most of the additional European uptake.

'***The manuscript was improved from the first round of reviews, especially in terms of highlighting the fact that there are not sufficient measurements to prove or disprove either theory. Despite this, the authors have chosen to ignore the recommendations to change the title. Given the evidence presented in the current manuscript, publishing it under this title is misleading. Perhaps a more realistic: "Elevated uptake of CO2 over Europe inferred from GOSAT XCO2 retrievals: the jury's still out". Given the lack of convincing evidence to support the conjecture contained in this study, I am very hesitant to recommend this study for publication in ACP.***

After the first round of review some of the conclusions were (rightly) scaled back, and some more nuance was added to the discussion. In the response to reviewers the authors suggest that they wanted to use this lack of evidence to prove or disprove either hypothesis as an impetus to improve our validation capacity. If the study were recast in this light, perhaps with suggestions about where these validation measurements could be implemented for maximum impact, it would be a more constructive contribution to the ongoing debate,

As stated above the main aims of this manuscript are to describe another possible explanation on the elevated European uptake inferred from GOSAT XCO2 retrievals, and to emphasize that more research is required to conclude robustly whether previous reports of elevated uptake of CO2 over Europe describe a real phenomenon. This study highlights the need for validating XCO2 under more and different observation conditions (such as tropical lands), and at different temporal and spatial scales (such as the monthly biases over Eastern and Western Europe). Also it highlights the need for developing and carefully testing inversion systems using more observations (as discussions in Appendix A and the inclusion of INV_ACOS_INS and INV_UOL_INS).

Minor points:

L25: observation without s
Change made.

L77: remove first comma, add "the" before global
Change made

L84 and 89: model -> modelling Map of sub-regions would be nice to see, especially with bias pattern...

That is a good suggestion, and we'd like to present them in the future study based on the latest GOSAT XCO2 retrievals.

L167: should be Figure 3? Why compare CONTRAIL under 3 km and HIPPO under 5 km? How sensitive is the analysis to these cut-offs?

Thanks for pointing out the typo, and we have changed it to Figure 3. We have used a lower cut-off (3km), just because it is easier for us to find the taking off/landing measurements around certain airports. A quick check shows no significant difference when the cut-off is increase to 5000m.

Figure 3: I really only see convincing evidence for the better agreement of the in-situ/TCCON inversion for February for Amsterdam, and maybe the autumn for Moscow. It's hard to say how significant this is, as none of the measurements are presented with error bars, or even standard deviations. This should be fixed. The broken magenta line should be removed: mixing the fluxes from two inversions leads to a meaningless result. Mixing concentration values from different simulations in an inversion (as is done later) makes more sense, as at least the final result is consistent with the input information. This, on the other hand, is like changing the boundary conditions *after* performing a regional inversion, and then discussing the results. It also adds little to the discussion, in my opinion.

The main purpose of Figure 3 is mainly to demonstrate that fluxes inferred from GOSAT XCO2 data have a higher CO2 inflow into the Europe. Because of the mass balance, overestimated flux inflows can result in overestimation of European uptake.

L205: Should this be Figure 3?

Thanks. We correct the typo

L228-229: Sentence should be rewritten.

We have change the sentence.

L234-235: mover "errors" before the parentheses.

L243: a few -> are few

Change made

L244: have -> has
Change made

L245: This sentence is redundant, isn't it? Given L228-229?

L247: are -> is
Change made

L264 & L265: "the" before "JPL ACOS team", "University of Leicester"

Change made

Review Report 4

The authors have addressed my previous comments. The only additional points are the following two typos that should be corrected:

Abstract, line 8: remove one of the duplicate "this"s

Line 167: The old Fig 2 is now the new Fig 3. -- change in text.

We'd like to thank Dr. David Baker for careful and constructive reviews. The typos have been corrected in the manuscript

Estimates of European uptake of CO_2 inferred from GOSAT X_{CO2} retrievals: sensitivity to measurement bias within and outwith Europe

Liang Feng[1], Paul I. Palmer[1], Robert Parker[2], Nicholas M. Deutscher[3,4], Dietrich G. Feist[5], Rigel Kivi[6], Isamu Morino[7], Ralf Sussmann[8]

1. National Centre for Earth Observation, School of GeoSciences, University of Edinburgh, UK.

2. National Centre for Earth Observation, Department of Physics and Astronomy, University of Leicester, UK.

3. Institute of Environmental Physics, University of Bremen, Germany.

4. Centre for Atmospheric Chemistry, University of Wollongong, Australia.

5. Max Planck Institute for Biogeochemistry, Jena, Germany.

6. FMI-Arctic Research Center, Sodankylä, Finland.

7. National Institute for Environmental Studies (NIES), Tsukuba, Japan.

8. Institute of Meteorology and Climate Research - Atmospheric Environmental Research KIT/IMK-IFU, Garmisch-Partenkirchen, Germany.

Estimates of the natural CO_2 flux over Europe inferred from in situ measurements of atmospheric CO_2 mole fraction have been used previously to check top-down flux estimates inferred from space-borne dry-air CO_2 column (X_{CO2}) retrievals. Several recent studies have shown that CO_2 fluxes inferred from X_{CO2} data from the Japanese Greenhouse gases Observing SATellite (GOSAT) and the Scanning Imaging Absorption Spectrometer for Atmospheric CHartographY (SCIAMACHY) have larger seasonal amplitudes and a more negative annual net CO_2 balance than those inferred from the in situ data. The cause of this elevated European uptake of CO_2 is still unclear, but recent studies have suggested that this this is a genuine scientific phenomenon. Here, we put forward an alternative hypothesis and show that realistic levels of bias in GOSAT data can result in an erroneous estimate of elevated uptake over Europe. We use a global flux inversion system to examine the relationship between measurement biases and estimates of CO_2 uptake from Europe. We establish a reference in situ inversion that uses an Ensemble Kalman Filter (EnKF) to assimilate conventional surface mole fraction observations and X_{CO2} retrievals from the surface-based Total Carbon Column Observing Network (TCCON). We use the same EnKF system to assimilate two independent versions of GOSAT X_{CO2} data. We find that the GOSAT-inferred European terrestrial biosphere uptake peaks during the summer, similar to the reference inversion, but the net annual flux is 1.40±0.19 GtC/a compared to a value of 0.58±0.14 GtC/a for our control inversion that uses only in situ data. To reconcile these two estimates, we perform a series of numerical experiments that assimilate observations with added biases or assimilate synthetic observations for which part or all of the GOSAT XCO_2 data are replaced with model data. We find that for our global flux inversions, a large portion (60-90%) of the elevated European uptake inferred from GOSAT data in 2010 is due to retrievals outside the immediate European region, while the remainder can largely be explained by a sub-ppm retrieval bias over Europe. We use a data assimilation approach to estimate monthly GOSAT X_{CO2} biases from the joint assimilation of in situ observations and GOSAT X_{CO2} retrievals. The inferred biases represent an estimate of systematic differences between GOSAT XCO2 retrievals and the inversion system at regional or sub-regional scales. We find that a monthly varying bias of up to 0.5 ppm can explain an overestimate of the annual sink of up to 0.20 GtC/a. Our results highlight the sensitivity of CO_2 flux estimates to regional observation biases, which have not been fully characterized by the current observation network. Without further dedicated measurements we cannot disprove that European ecosystems are taking up a larger-than-expected amount of CO_2. More robust inversion systems are also needed to infer consistent fluxes from multiple observation types.

1. Introduction

Observed atmospheric variations of carbon dioxide (CO_2) are due to atmospheric transport and surface flux processes. Using prior knowledge of the spatial and temporal distribution of these fluxes and atmospheric transport it is possible to infer (or invert for) the a posteriori estimate of surface fluxes from atmospheric concentration data. The geographical scarcity of such observations precludes robust flux estimates for some regions due to large uncertainties associated with meteorology and a priori fluxes. Arguably, our knowledge of top-down estimates of regional CO_2 fluxes, particularly at tropical and high northern latitudes, have not significantly improved for over a decade [Gurney et al., 2002; Peylin et al., 2013], reflecting the difficulty of maintaining a surface measurement programme over vulnerable and inhospitable ecosystems. Atmospheric transport model errors compound errors introduced by poor observation coverage, resulting in significant differences between flux estimates on spatial scales < O(10,000 km) [e.g. Law et al., 2003; Yuen et al., 2005; Stephens et al., 2007]

The Greenhouse gases Observing SATellite (GOSAT), a space-borne mission launched in a sun-synchronous orbit in early 2009, was purposefully designed to measure CO_2 columns using short-wave IR wavelengths. Validation of current X_{CO2} column retrievals using co-located upward-looking FTS measurements of the Total Carbon Column Observing Network (TCCON) [Wunch et al., 2011] show a standard deviation of 1.6-2.0 ppm (e.g., Parker et al., 2012). Their global biases are typically smaller than 0.5 ppm [Oshchepkov et al., 2013]. The disadvantage of using the TCCON is that sites are mainly at northern extra-tropical latitudes with little or no coverage where our knowledge of the carbon cycle is weakest. Many surface flux estimation algorithms are particularly sensitive to systematic errors so that sub-ppm biases can still significantly change the patterns of regional flux estimates [Chevallier et al., 2010]. This is further complicated by the seasonal coverage of GOSAT data at high latitudes during winter months when solar zenith angles are too large to retrieve reliable values for X_{CO2} [Liu et al., 2014].

Several independent studies have shown that regional flux distributions inferred from GOSAT X_{CO2} retrievals are significantly different from those inferred from in situ data [Basu et al., 2013; Deng et al., 2013; Chevallier et al., 2014]. In particular, these studies report a larger-than-expected annual net emission over tropical continents and a larger-than-expected net annual uptake over Europe. While the GOSAT inversions suffer from larger observation errors, atmospheric transport errors and issues from the seasonal coverage of higher latitudes, the in situ inversions are also unreliable over many regions due to poor coverage and atmospheric transport errors. Inter-comparisons revealed significant inconsistency in regional flux estimates inferred from in-situ observations by using different inversion systems, over many regions important for global carbon cycle, including Europe

[Peylin et al., 2013]. Consequently, there is an ongoing debate about whether a recent study that shows a large European uptake of CO_2 [Reuter et al., 2014] reflects a real phenomenon or is an artefact due to deficiencies both in the observations and in the inverse modelling.

We report the results from a small set of experiments that show systematic bias can introduce a large difference between European fluxes inferred from GOSAT and those inferred from in situ data by using a global flux inversion approach. In the next section we provide an overview of the inverse model framework used to interpret data from the in-situ observation network (including both the conventional surface observation network and the relatively new TCCON network), and from the space-based GOSAT X_{CO2} data. In section 3, we present results from two groups of global inversion experiments that characterize the role of systematic bias in regional flux estimates. Further experiments for quasi-regional flux inversions are presented in Appendix A. In section 4, we use a modified version of the inverse model framework to estimate monthly biases by jointly assimilating all data. We conclude the paper in section 5.

2. Description and Evaluation of Control In-situ and GOSAT Experiments

We use the GEOS-Chem global chemistry transport model to relate surface fluxes to the observed variations of atmospheric CO_2 concentrations [Feng et al., 2009] at a horizontal resolution of 4^ox5^o, driven by GEOS-5 meteorological analyses from the Global Modeling and Assimilation Office Global Circulation Model based at NASA Goddard Space Flight Centre. We use an Ensemble Kalman Filter (EnKF) [Feng et al., 2009; 2011] to estimate regional fluxes from in situ or GOSAT observations for three years from 2009-2011, but we focus on 2010 to minimize error due to spin-up and edge effects. We estimate monthly fluxes on a spatial distribution that is based on TransCom-3 [Gurney et al., 2002] with each continental region further divided equally into 12 sub-regions and each ocean region further divided equally into 6 sub-regions. As a result, we estimate fluxes for 199 regions, compared to 144 regions we have used in previous studies [Feng et al, 2009; Chevallier et al., 2014].

In all global inversion experiments we assume the same set of a priori flux inventories, including: (1) monthly fossil fuel emissions [Oda and Maksyutov, 2011]; (2) weekly biomass burning emissions (GFED v3.0) [van der Werf et al., 2010]; (3) monthly oceanic surface CO_2 fluxes [Takahashi et al., 2009]; and (4) 3-hourly terrestrial biosphere-atmosphere CO_2 exchange [Olsen and Randerson, 2004]. We assume that the a priori uncertainty for each land sub-region is proportional to a combination of the net biospheric emission (70%) at the current month, and its annual variation (30%). We also assume that the a priori errors are correlated with each other with a spatial correlation length of 800 km, and a temporal correlation of 1 month [Chevallier et al., 2014]. We then determine the coefficient for the assumed a priori uncertainty by scaling the aggregated annual uncertainty over all 133 land sub-regions to 1.9 GtC/a. In particular, the resulting annual a priori uncertainty for European region is about 0.52 GtC/a, with the monthly uncertainty varying from 2.0 GtC/a for the summer months to about 0.8 GtC/a for winter months, which is generally larger than the a priori monthly uncertainty used by Deng et al. (2014). Prior uncertainties over oceans are determined under similar assumption but with a longer spatial correlation (1500 km), and a smaller aggregated annual error (0.6 Gt/a). Our experiments show that doubling the a priori uncertainty increases the European uptake inferred from GOSAT data by about 0.21 GtC/a (from 1.40 GtC/a to 1.61 GtC/a), compared to a smaller increase of 0.09 GtC/a for the in situ inversion (from 0.58 GtC/a to 0.67 GtC/a).

Our control inversion experiment (INV_TCCON, Table 1 and Figure 1) assimilates in situ observations, including the conventional surface observations at 76 sites [Feng et al., 2011] and, in particular, the total column X_{CO2} retrievals from all the TCCON sites of the GGG2014 dataset (see Wennberg et al., 2014, and https://tccon-wiki.caltech.edu for more details) to improve observation constraints. We use daytime (09:00 to 15:00 local time) mean TCCON retrievals, with the observation errors determined by the standard deviation about their daytime mean. To account for the inter-site biases as well as the model representation errors, we enlarge the TCCON observation errors by 0.5 ppm. Including TCCON observations increases the annual net uptake over Europe in 2010 from 0.49 GtC/a, as inferred from surface observations only, to 0.58 GtC/a. The increase is mainly due to a larger summer uptake. TCCON data also reduce the a posteriori uncertainty by about 15% from 0.16 GtC/a to 0.14 Gt/a. 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].

For the two control GOSAT inversions (Figure 1), we use two independent data sets: (1) X_{CO2} retrievals from JPL ACOS team (v3.3) [Osterman et al., 2013] (INV_ACOS); and (2) the full-physics X_{CO2} retrievals (v4.0) from the University of Leicester [Cogan et al., 2012] (INV_UOL). For both data sets, we assimilate only the H-gain data over land regions, and apply the bias corrections recommended by the data providers. We double the reported observation errors, as suggested by the retrieval groups.

As a performance indicator for our ability to fit fluxes to observed X_{CO2} concentrations, we compare a posteriori model concentrations with GOSAT X_{CO2} retrievals and show that INV_ACOS and INV_UOL agree much better than INV_TCCON. For example, the bias against ACOS X_{CO2} retrievals is -0.45 ppm for INV_TCCON and 0.02 ppm for INV_ACOS with a corresponding reduction in the global standard deviation from 1.69 ppm to 1.57 ppm. However comparison of GOSAT a posteriori concentrations against independent HIPPO-3 measurements is worse than INV_TCCON with a positive bias of 0.47 ppm and 0.66 ppm for INV_ACOS and INV_UOL, respectively, which are mainly caused by the overestimation of CO_2 concentrations (~1.5-2.0 ppm) at low latitudes (Figure 2).

3. Results

Figure 1 and Table 1 shows the three inversion experiments, INV_TCCON, INV_ACOS, and INV_UOL, have similar European uptake values in June 2010 (0.69 GtC/m for INV_TCCON and ~0.72 GtC/m for GOSAT inversions), and are generally consistent with other GOSAT inversion experiments (e.g., Deng et al., 2014; Chevallier et al., 2014). But the GOSAT inversions have an annual net uptake of about 1.40±0.19 GtC/a compared to the in situ inversion of 0.58±0.14 GtC/a. Figure 1 also shows significant differences between their monthly flux estimates in early spring and winter when there is only sparse GOSAT observation coverage, particularly over northern Europe. Both INV_UOL and INV_ACOS have a cumulative total of about 0.51 GtC more uptake than INV_TCCON during February-April of 2010, with a further 0.37 GtC uptake accumulated over the following summer and autumn. This larger uptake is partially cancelled out by larger emissions (0.17-0.08 GtC) at the end of 2010.

Figure 2 shows that INV_TCCON a posteriori CO_2 mole fractions agree well with the independent HIAPER Pole-to-Pole Observations (HIPPO-3) aircraft measurements below 5 km over the Pacific

Ocean in 2010 [Wofsy et al., 2010], with a small bias of 0.05 ppm, and a sub-ppm standard deviation of 0.87 ppm. Figure 3 shows further evaluation of a posteriori CO_2 mole fractions using descending and ascending profile observations over two European airports from the CONTRAIL experiment [Machida et al., 2008]. We calculate monthly mean CONTRAIL measurements during 2010 using data below 3 km, where there is greater sensitivity to local surface fluxes. Our current model resolution precludes small-scale sources (or sinks) so we expect model bias. We find that INV_TCCON agrees best with CONTRAIL observations, in particular at the beginning of the 2010, partially reflecting the poor GOSAT X_{CO2} coverage over Europe during the winter and early spring. However, we cannot conclude from the slightly degraded agreement with CONTRAIL (as well as with HIPPO-3) that the European uptake inferred from GOSAT data is incorrect, because unaccounted small local emissions/sinks, and model transport errors can affect the comparison against aircraft observations.

Figure 3 also presents an additional model simulation forced by a hybrid flux (denoted by the magenta broken line) where the INV_TCCON a posteriori fluxes outside Europe are replaced by the results from INV_ACOS. The resulting CO_2 concentrations from these hybrid fluxes are, as expected, higher than the a posteriori model concentrations for INV_ACOS because of the larger European emissions (i.e., less uptake) inferred by INV_TCCON. But they are also systematically higher than the INV_TCCON simulation, in particular during spring months, despite the same European fluxes being used to force these two simulations. This suggests an overestimate of CO_2 transported into the European region by the GOSAT inversions. Further comparison of the INV_TCCON simulation and the hybrid run reveals that systematic differences in the inflow into the European domain can affect the atmospheric X_{CO2} gradient across this region. In the INV_TCCON simulation, the mean X_{CO2} difference between east (east of 20^{o}E) and west (west of 20^{o}E) Europe is ~0.04 ppm for May, 2010, which is increased to 0.16 ppm in the hybrid run (cf. E-W X_{CO2} gradient of -0.20 ppm for GOSAT ACOS data).

To understand the differences between the INV_TCCON and GOSAT inversions, we conducted two groups of sensitivity tests (Table 1 and Figure 4). First, we replaced all or part of the GOSAT X_{CO2} retrievals assimilated in INV_ACOS with those from a model simulation forced by the a posteriori fluxes from INV_TCCON. In experiment INV_ACOS_MOD_ALL (Figure 4), where we replace all GOSAT data with CO_2 concentrations inferred from INV_TCCON, we reproduce INV_TCCON with small exceptions at beginning of 2010, reflecting the seasonal variation in GOSAT coverage. In a related experiment INV_ACOS_MOD_NOEU for which we replace X_{CO2} retrievals outside Europe with the model simulation, the differences between the GOSAT and in situ inversions are significantly reduced, particularly over the period with limited observation coverage, although the actual X_{CO2} retrievals are still assimilated over Europe. The simulated GOSAT data outside Europe reduces the estimate of European uptake from 1.40 GtC/a to 0.88 GtC/a. In other words, the GOSAT observations outside the European region are responsible for about 60% (0.52 GtC/a) of the total enhanced European sink (0.82 GtC/a) with the remainder (0.30 GtC/a) due to observations taken directly over Europe. For INV_UOL, when we replace the X_{CO2} data outside Europe by the a posteriori INV_TCCON model simulations, European uptake is reduced to 0.67 GtC/a (INV_UOL_MOD_NOEU, Table 1), indicating an external contribution of nearly 90% to the enhanced uptake of 0.82 GtC/a. Together with Figure 3, these results suggest that GOSAT inversions result in an overestimated CO_2 inflow. This will subsequently lead to the fitted European flux having to compensate, via mass balance, by being erroneously low even when un-biased GOSAT X_{CO2} data are assimilated over the immediate European region. We find similar effects in the quasi-regional inversions (Figure A1), where only observations within European region are assimilated, with flux estimates from INV_TCCON or from INV_ACOS being used to provide lateral boundary conditions around Europe.

Second, we crudely demonstrate how regional bias could explain the remaining discrepancy of up to 0.30 GtC/a between GOSAT and in situ inversions over Europe. In our experiment INV_ACOS_SPR_0.5ppm, we add a bias of +0.5 ppm to the GOSAT ACOS retrievals within Europe taken in February-April, inclusively, which effectively reduces the uptake by 0.1 GtC/a from 1.40 GtC/a to 1.30 GtC/a. Similarly, when the bias of +0.5 ppm is added to the GOSAT data taken in June-August we find a larger reduction of 0.15 GtC/a for the summer peak uptake (INV_ACOS_SUM_0.5ppm), partially due to a larger a priori uncertainty and denser GOSAT coverage during the summer. These results emphasize the importance of characterizing sub-ppm regional bias to avoid erroneous flux estimates.

4. Bias estimation.

Here we demonstrate a simple approach to quantify systematic bias in X_{CO2} retrievals based on a simple on-line bias correction scheme. We assimilate the GOSAT X_{CO2} retrievals together with the surface and TCCON observations in two experiments: INV_ACOS_INS and INV_UOL_INS (Table 1). 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. To investigate the spatial pattern of the X_{CO2} biases within Europe, we split Europe into West Europe (west of 20°E) and East Europe (east of 20°E). We assume that a priori for monthly biases is 0.0±0.5 ppm. For simplicity, we have assumed the a priori errors for regional X_{CO2} biases are not correlated. Compared to the off-line comparisons between GOSAT X_{CO2} retrieval and model concentrations, the main advantage of the on-line bias estimation is that the uncertainties associated with error in flux estimates can be partially taken into account. However, biases derived by this approach reflect the systematic difference between the model simulation and GOSAT data over large (continental) regions, which also contain systematic model errors (such as the atmospheric transport and representation errors). In addition, the inversion results are affected by the relative weights assigned to different data sets, as well as by the relative prior uncertainty assumed for surface fluxes and for the observation bias. The seasonal variation of the mean CO_2 concentration is an important sign of the underlined biosphere seasonal cycle. We show in Appendix A that when we inflate the a priori uncertainty for the assumed observation bias, the observation constraints on flux estimate will become weaker. Also, the on-line bias correction is only effective for detecting and correcting bias at specified patterns, which may increase the sensitivity to other uncharacterized systematic errors. Despite these weaknesses, a joint data assimilation approach can exploit complementary constraints from in situ and satellite X_{CO2} data: for example there are few GOSAT observations over northern Europe during autumn and winter months, while Eastern Europe has few in situ observations. We have also limited the a priori uncertainty for the monthly observation biases to 0.5 ppm.

In the joint inversions INV_ACOS_INS and INV_UOL_INS, the annual European uptake is estimated to be 0.62 GtC/a, and 0.67 GtC/a, respectively (Table 1), which is close to the reference value of 0.58 GtC/a inferred from the in situ observations. To test the impact of the on-line bias correction, we set the a priori uncertainty of regional X_{CO2} bias to be 0.01 ppm so that on-line bias correction is effectively turned off. As a result, the annual European uptake for INV_ACOS_INS is increased by

0.15 GtC to 0.77 GtC/a, which is close to INV_ACOS_MOD_NOEU, but about 55% of the GOSAT only inversions (1.40 GtC/a).

Figure 5 shows the estimated monthly biases in ACOS and UOL X_{CO2} retrievals over East and West Europe during 2010. Monthly biases are typically smaller than 0.5 ppm over the two regions, but have different seasonal cycles. Additional experiment shows that after ACOS X_{CO2} data over Europe have been corrected for the inferred biases, the European annual uptake by INV_ACOS is reduced by 0.20 GtC/a, representing more than half of the contribution from GOSAT observations within Europe. This result is consistent with our sensitivity tests. The effect of bias correction is much smaller for INV_UOL (about 0.07 GtC/a), because of the different bias patterns. Differences in GOSAT X_{CO2} retrievals and their effects on regional flux estimates have also been investigated in previous studies (e.g., Takagi et al., 2014).

5. Discussion and Conclusions

We used an ensemble Kalman Filter to infer regional CO_2 fluxes from three different CO_2 data sets: 1) surface in situ mole fraction observations and TCCON X_{CO2} retrievals ; 2) GOSAT X_{CO2} retrievals from the JPL ACOS team; and 3) GOSAT X_{CO2} retrievals from the University of Leicester. Our results, consistent with previous studies, show that these GOSAT data in a global flux inversion context result in a significantly larger European uptake than inferred from in situ data during 2010.

We showed using sensitivity experiments that a large portion (60-90%) of the elevated European uptake of CO_2 is related to the systematically higher model CO_2 mass being transported into Europe, due to the assimilation of GOSAT X_{CO2} data outside the European region. We find some evidence using aircraft observations over the Pacific that GOSAT a posteriori fluxes result in higher CO_2 concentration over lower latitudes. But limited observation coverage and unaccounted model errors prevent us from confidently concluding that GOSAT X_{CO2} data are biased high or low. Our global and quasi-regional (Appendix A) flux inversion experiments show that the main consequence of the elevated CO_2 inflow to the European domain is that the European uptake must increase because of mass balance, even when GOSAT X_{CO2} retrievals within the European domain are not biased. A crude sensitivity test (INV_ACOS_OUT_0.5ppm) shows that reducing ACOS X_{CO2} data outside the European region by 0.5 ppm will reduce European annual uptake from 1.40 GtC/a to 0.98 GtC/a. Erroneous interpretation of X_{CO2} data can result from analyses if biased boundary conditions are not addressed. However, as shown in Appendix A, a gross mis-characterization and correction of bias may weaken observation constraints, which can also lead to erroneous flux estimates.

We also showed using sensitivity tests that sub-ppm bias can explain the remaining 0.30 GtC/a flux difference between the in situ inversion and INV_ACOS after accounting for biased boundary conditions. By simultaneously assimilating the in situ and GOSAT observations to estimate surface fluxes and monthly X_{CO2} biases, we infer a monthly observation bias that is typically less than 0.5 ppm over East and West Europe, but is able to cause an elevated sink of up to 0.20 GtC/a. The inferred monthly biases for UOL X_{CO2} are also not the same as the ACOS X_{CO2} data, particularly over West Europe during the summer months. This level of sensitivity of regional flux estimate to time-varying sub-ppm observation bias highlights the challenges we face as a community when evaluating X_{CO2} retrievals using current observation networks.

Flux estimates are sensitive to a priori assumptions, idiosyncrasies of applied inversion algorithms, and the underlying model atmospheric transport [Chevallier et al., 2014; Peylin et al., 2014; Reuter et al., 2014]. The possible presence of regional observation biases further complicates the inter-comparisons of flux estimates based on different inversion approaches, as they may have different sensitivities to certain observation biases. In our assimilation of ACOS X_{CO2} retrievals, we find that doubling the a priori flux error (INV_ACOS_DBL_ERR) increases the estimated European uptake from 1.40 GtC/a to 1.61 GtC/a, consistent with the hypothesis on the increased vulnerability to the observation biases both within and outside Europe when using weak a priori constraints. In contrast, doubling the a priori flux errors only increases the uptake by 0.05 GtC/a to 0.67 GtC/a for the joint data assimilation (INV_ACOS_INS_DBL_ERR), with very little changes in the estimated biases (not shown). Examples in Appendix A also demonstrate different responses to regional and sub-regional biases before and after an on-line scheme is used to correct the systematic error across Europe. These differences emphasize the need for a closer examination of the responses of the inversion systems to the assimilated observations, as well as to their possible biases, to help understand the inter-model variations in estimated regional fluxes.

Complicated interactions between observations and the assimilation system also mean that our present study does not exclude other possible causes for the elevated European uptake reported by previous research from assimilation of GOSAT data. Instead, it highlights the adverse effects of possibly uncharacterized regional biases in current GOSAT X_{CO2} retrievals that can attract erroneous interpretation of resulting regional flux estimates. A more thorough evaluation of the X_{CO2} retrievals using independent and sufficiently accurate/precise observations is urgently required to increase the confidence of regional CO_2 flux estimates inferred from space-based observations. Without additional observations, we cannot rule out either the lower European uptake estimate of around 0.6 GtC/a (inferred from the in situ inversion INV_TCCON and the joint inversion INV_ACOS_INS and INV_UOL_INS) or the higher European uptake estimate of around 1.40 GtC/a (inferred from GOSAT data). There is also no sufficient reason to believe that the mean value among these diverse estimates is more reliable, because our study suggests that small systematic errors can result in significant differences in the estimated fluxes, and the influences of random errors have also not been fully quantified. The observational density required to infer flux estimates over a limited spatial domain such as Europe is crucial. For the time frame of this analysis, the TCCON network provided good coverage for Europe, North America, South East Asia and Australia/New Zealand. Great efforts were also taken to reduce inter-station biases. In future the TCCON measurement network may be supported by smaller, more mobile FTIR instruments, which can be established, at least on a campaign basis, in tropical and high latitude locations where observational gaps are greatest.

Our joint data assimilation approach assimilates in situ and space-borne observations. They also provide estimates of systematic differences between XCO2 retrievals and the inversion system at regional and sub-regional scales. However the resulting differences will include the observation biases and deficiencies in the underlined inversion approaches. To achieve consistent flux estimates inferred from assimilating multiple data sets using different inversion approaches, we need to better quantify observation and model errors, and need to better understand the sensitivity of each inversion system to the assimilated observations as well as to their possible biases. It is difficult to develop a robust bias correction scheme before properly characterizing observation biases and the responses by the inversion system.

Acknowledgements

LF and PIP designed the experiments and wrote the paper, RP provided the GOSAT XCO2 data and comments on the paper, and NMD, DGF, RK, IM, and RS provided access to TCCON XCO2 data and comments on the paper. Work at the University of Edinburgh was partly funded by the NERC National Centre for Earth Observation (NCEO). PIP gratefully acknowledges funding from the NCEO and his Royal Society Wolfson Research Merit Award. Work at the University of Leicester was funded by NCEO and the European Space Agency Climate Change Initiative (ESA-CCI). The TCCON Network is supported by NASA's Carbon Cycle Science Program through a grant to the California Institute of Technology. The TCCON stations from Bialystok, Orleans and Bremen are supported by the EU projects InGOS and ICOS-INWIRE, and by the Senate of Bremen. TCCON measurements at Eureka were made by the Canadian Network for Detection of Atmospheric Composition Change (CANDAC) with additional support from the Canadian Space Agency. The authors thank the NASA JPL ACOS team for providing their X_{CO2} retrievals. We also thank the CONTRAIL and HIPPO team for their observations used in our validations. We thank G. J. Collatz and S. R. Kawa for providing NASA Carbon Monitoring System Land Surface Carbon Flux Products: http://nacp-files.nacarbon.org/nacp-kawa-01/. We are grateful to Dr. Hartmut Bösch, Dr. Chris O'Dell and Dr. Thorsten Warneke for their helpful comments on the manuscript.

References

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO_2 fluxes estimate from GOSAT retrievals of total column CO_2 , Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.

Chevallier, F., Feng, L., Bösch, H., Palmer, P. I., and Rayner, P.: On the impact of transport model errors for the estimation of CO_2 surface fluxes from GOSAT observations, Geophys. Res. Lett., 37, L21803, doi:10.1029/2010GL044652, 2010.

Chevallier, F., Palmer, P. I., Feng, L., Bösch, H., O'Dell, C., and Bousquet, P.: Towards robust and consistent regional CO_2 flux estimates from in situ and space-borne measurements of atmospheric CO_2, Geophys. Res. Lett., 41, 1065–1070, doi:10.1002/2013GL058772, 2014.

Cogan, A. J., Boesch, H., Parker, R. J., Feng, L., Palmer, P. I., L Blavier, J.-F., Deutscher, N. M., Macatangay, R., Notholt, J., Roehl, C., Warneke, T., and Wunch, D.: Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite: Comparison with ground-based TCCON observations and GEOS-Chem model calculations, J. Geophys. Res., 117, D21301, doi:10.1029/2012JD018087, 2012.

Deng, F., Jones, D. B. A., Henze, D. K., Bousserez, N., Bowman, K. W., Fisher, J. B., Nassar, R., O'Dell, C., Wunch, D., Wennberg, P. O., Kort, E. A., Wofsy, S. C., Blumenstock, T., Deutscher, N. M., Griffith, D. W. T., Hase, F., Heikkinen, P., Sherlock, V., Strong, K., Sussmann, R., and Warneke, T.: Inferring regional sources and sinks of atmospheric CO_2 from GOSAT XCO_2 data, Atmos. Chem. Phys., 14, 3703-3727, doi:10.5194/acp-14-3703-2014, 2014.

Feng, L., Palmer, P. I., Bösch, H., and Dance, S.: Estimating surface CO2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman Filter, Atmos. Chem. Phys., 9, 2619–2633, doi:10.5194/acp-9-2619-2009, 2009.

Feng, L., Palmer, P. I., Yang, Y., Yantosca, R. M., Kawa, S. R., Paris, J.-D., Matsueda, H., and Machida, T.: Evaluating a 3-D transport model of atmospheric CO2 using ground-based, aircraft, and space-borne data, Atmos. Chem. Phys., 11, 2789–2803, doi:10.5194/acp-11-2789-2011, 2011.

Fraser, A., Palmer, P. I., Feng, L., Bösch, H., Parker, R., Dlugokencky, E. J., Krummel, P. B., and Langenfelds, R. L.: Estimating regional fluxes of CO_2 and CH_4 using space-borne observations of XCH_4: XCO_2, Atmos. Chem. Phys., 14, 12883-12895, doi:10.5194/acp-14-12883-2014, 2014.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L.,Chen, Y., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, doi: 10.1038/415626a, 2002.

Liu, J., Bowman, K. W., Lee, M., Henze, D. K., Bousserez, N., Brix, H., Collatz, G. J., Menemenlis, D., Ott, L., Pawson, S., Jones, D., and Nassar, R.: Carbon monitoring system flux estimation and attribution: impact of ACOS-GOSAT XCO2 sampling on the inference of terrestrial biospheric sources and sinks, Tellus B, 66, 22486, doi: 10.3402/tellusb.v66.22486, 2014.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., Goto, K., Nakazawa, T., Ishikawa, K., and Ogawa, T.: Worldwide measurements of atmospheric CO_2 and other trace gas species using commercial airlines, J. Atmos. Ocean. Tech., 25, 1744–1754, doi:10.1175/2008jtecha1082.1, 2008.

Oda, T., and Maksyutov, S.: A very high-resolution (1 km×1 km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of nighttime lights, Atmos. Chem. Phys., 11, 543–556, doi:10.5194/acp-11-543-2011, 2011.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res., 109, D02301, doi:10.1029/2003JD003968, 2004.

Oshchepkov, S., Bril, A., Yokota, T., Wennberg, P. O., Deutscher, N. M., Wunch, D., Toon, G. C., Yoshida, Y., O'Dell, C. W., Crisp, D., Miller, C. E., Frankenberg, C., Butz, A., Aben, I., Guerlet, S., Hasekamp, O., Boesch, H., Cogan, A., Parker, R., Grith, D., Macatangay, R., Notholt., J., Sussmann, R., Rettinger, M., Sherlock, V., Robinson., J., Kyrö, E., Heikkinen, P., Feist, D. G., Morino, I., Kadygrov, N., Belikov, D., Maksyutov, S., Matsunaga, T., Uchino, O., and Watanabe, H.: Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space. Part 2: Algorithm intercomparison in the GOSAT data processing for CO_2 retrievals over TCCON sites, J. Geophys. Res.-Atmos., 118, 1493–1512, doi:10.1002/jgrd.50146, 2013.

Osterman, G., Eldering, A., Avis, C., O'Dell, C., Martinez, E., Crisp, D., Frankenberg, C., Fisher, B., and Wunch, D.: ACOS level 2 standard product data user's guide, v3.3, available at: http://oco.jpl.nasa.gov/files/oco/ACOS_v3.3_DataUsersGuide.pdf, 2013.

Parker, R. and the GHG-CCI project team: Product user guide (PUG) for the University of Leicester full-physics XCO2 GOSAT data product (CO2_GOS_OCFP) for the GHG-CCI project of ESA's climate change initiative, 15 pp., version 1, 12 August 2013, available at: http://www.esa-ghg-cci.org/sites/default/files/documents/public/documents/PUG_GHG-CCI_CO2_GOS_OCFP_final.pdf, 2013.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.

Prather, M. J., Zhu, X., Strahan, S. E., Steenrod, S. D., and Rodriguez, J. M.: Quantifying errors in trace species transport modeling, P. Natl. Acad. Sci. USA, 105, 19617–19621, doi:10.1073/pnas.0806541106, 2008.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi, R., Sussmann, R., Machida, T., Matsueda, H., and Sawa, Y.: Satellite-inferred European carbon sink larger than expected, Atmos. Chem. Phys. Discuss., 14, 21829–21863, doi:10.5194/acpd-14-21829-2014, 2014.

Stephens, B. B., Gurney, K.R., Tans, P. P., Sweeney, C., Peters, W. , Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R.L., Steele, L.P., Francey, R. J., Denning, A. S.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, Science, 316, 1732–1735, doi: 10.1126/science.1137004, 2007.

Takagi, H., Houweling, S., Andres, R. J., Belikov, D., Bril, A., Boesch, H., Butz, A., Guerlet, S., Hasekamp, O., Maksyutov, S., Morino, I., Oda, T., O'Dell, C. W., Oshchepkov, S., Parker, R., Saito, M., Uchino, O., Yokota, T., Yoshida, Y., and Valsala, V.: Influence of differences in current GOSAT XCO2 retrievals on surface flux estimation, Geophys. Res. Lett., 41, 2598–2605, doi:10.1002/2013GL059174, 2014.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C. E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinho, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C. S., Delille, B., Bates, N. R., and de Baar, H. J. W.: Climatological mean and decadal changes in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, Deep-Sea Res. Pt. II, 56, 554–577, doi:10.1016/j.dsr2.2008.12.009, 2009.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmos. Chem. Phys., 10, 11707–11735, doi:10.5194/acp-10-11707-2010, 2010.

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeo-sciences, 10, 1983–2000, doi:10.5194/bg-10-1983-2013, 2013.

Wennberg, P. O., Roehl, C., Wunch, D., Toon, G. C. , Blavier, J.-F. , Washenfelder, R. , Keppel-Aleks, G., Allen, N. , Ayers, J: TCCON data from Park Falls, Wisconsin, USA, Release GGG2014R0. TCCON data archive, hosted by the Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A. http://dx.doi.org/10.14291/tccon.ggg2014.parkfalls01.R0/1149161, 2014.

Wofsy, S. C., The HIPPO Science Team, and Cooperating Modellers and Satellite Teams: HIAPER pole-to-pole observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, P. R. Soc. A, 369, 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O.: The total carbon column observing network, Philos. T. R. Soc. A, 369, 2087–2112, doi:10.1098/rsta.2010.0240, 2011.

Tables

Name	Data	Flux (GtC/a)	Uncertainty (GtC/a)
INV_TCCON	In-situ Flask and TCCON X_{CO2}	-0.58	0.14
INV_ACOS	ACOS X_{CO2} retrievals	-1.40	0.19
INV_UOL	UOL X_{CO2} retrievals	-1.4	0.20
INV_ACOS_MOD_ALL	Model simulation of ACOS X_{CO2} by using INV_TCCON posterior fluxes	-0.64	0.19
INV_ACOS_MOD_NOEU	As INV_ACOS_MOD_ALL but the real ACOS X_{CO2} retrievals are assimilated within Europe.	-0.88	0.19
INV_UOL_MOD_NOEU	As INV_UOL, but outside the Europe, UOL X_{CO2} retrievals are replaced with INV_TCCON simulations.	-0.67	0.19
INV_ACOS_MOD_ONLYEU	As INV_ACOS, but X_{CO2} retrievals within EU are replaced by INV_TCCON simulations	-1.17	0.19
INV_ACOS_OUT_0.5ppm	As INV_ACOS, but a bias of -0.5 ppm has been added to X_{CO2} retrievals outside Europe.	-0.98	0.19
INV_ACOS_SPR_0.5ppm	As INV_ACOS, but 0.5 ppm bias has been added to the European data in February, March, and April.	-1.30	0.19
INV_ACOS_SUM_0.5ppm	As INV_ACOS, but 0.5 ppm bias has been added to the European data in June, July, and August.	-1.25	0.19
INV_ACOS_INS	ACOS X_{CO2} retrievals and In-situ flask and TCCON data	-0.62	0.13
INV_UOL_INS	UOL X_{CO2} retrievals and in-situ flask and TCCON data	-0.67	0.13
INV_ACOS_DBL_ERR	ACOS X_{CO2} retrievals, but the a priori uncertainties have been doubled	-1.61	0.27
INV_ACOS_INS_DBL_ERR	GOSAT ACOS X_{CO2} retrievals and In-situ flask and TCCON data but the a priori flux uncertainties have been doubled	-0.67	0.16

Table 1: The magnitude and uncertainty of the European annual CO_2 biosphere flux (GtC/a) from 14 global flux inversion experiments. 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.

Figures

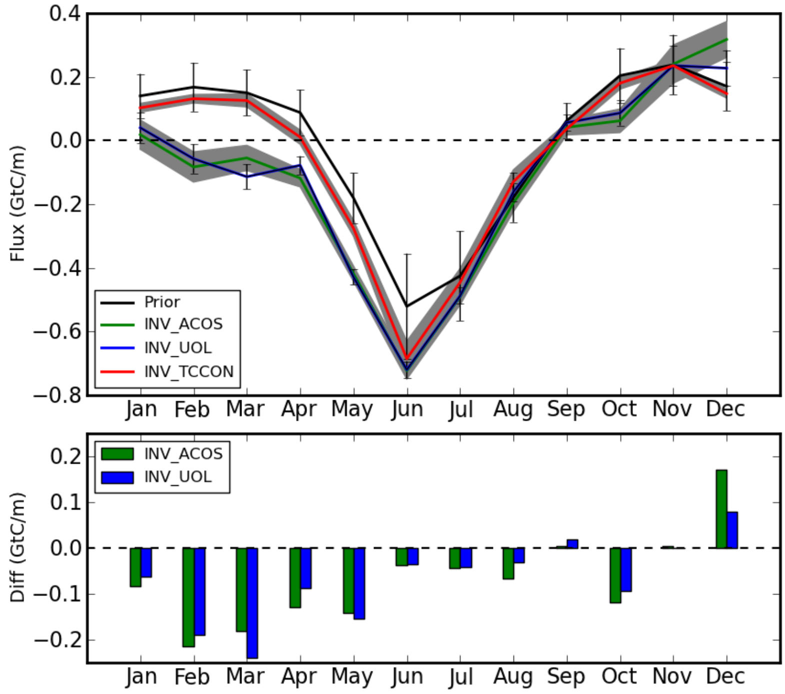

Figure 1. Monthly a posteriori estimates (GtC/m) for European biospheric CO_2 fluxes in 2010 using three inversion experiments (top panel): 1) INV_TCCON (red line), 2) INV_ACOS (green line), and INV_UOL (blue line). The black line denotes a priori values. The vertical black lines and grey shading denotes the uncertainties of the corresponding a priori or a posteriori flux estimates, respectively. Differences in monthly CO_2 uptake (GtC/m) between INV_TCCON and two GOSAT inversions (bottom panel): INV_ACOS (green bars) and INV_UOL (blue bars).

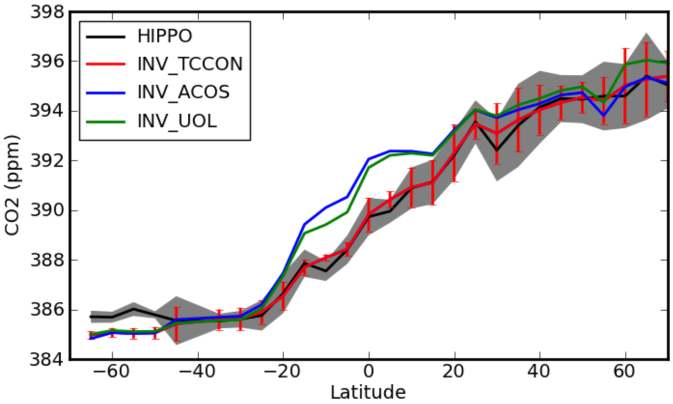

Figure 2: HIPPO-3 and GEOS-Chem model atmospheric CO2 mole fractions (ppm) over the Pacific Ocean below 5 km (black). GEOS-Chem is driven by different a posteriori flux estimates: 1) INV_TCCON (red), 2) INV_ACOS (blue), and 3) INV_UOL (green). HIPPO-3 and model CO_2 mole fractions are binned into 5^{o} latitude boxes. We calculate the mass-weighted average over these latitude boxes by assigning each HIPPO-3 and GEOS-Chem model value a weighting factor according to the observation altitude (air pressure). The grey envelope (red vertical lines) indicates the one standard deviation of HIPPO-3 measurements (INV_TCCON model values) within each latitude box.

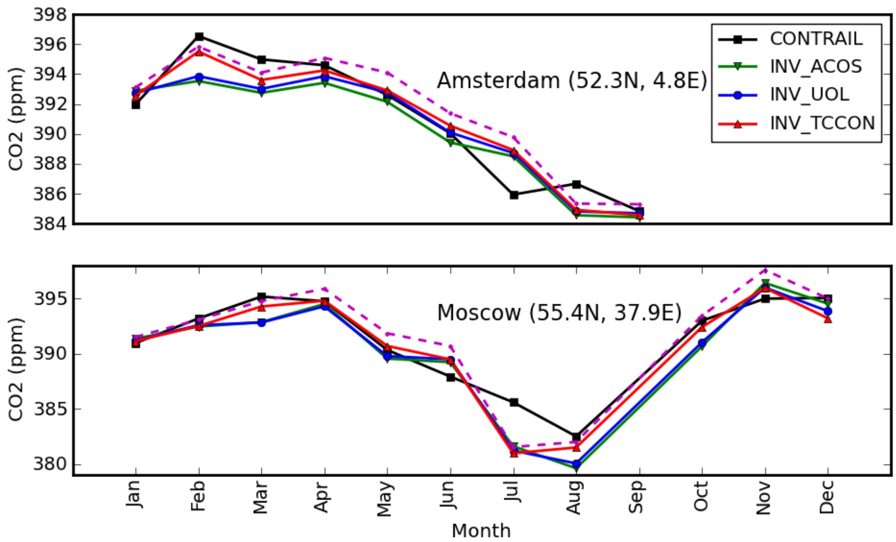

Figure 3: Monthly mean observed and model a posteriori model CO_2 mole fractions (ppm) below 3km above Amsterdam (the top panel) and Moscow (the bottom panel) airports during 2010, respectively [Machida et al., 2008]. The three sets of a posteriori model concentrations are inferred from three inversion experiments: INV_TCCON (red line), INV_ACOS (green line), and INV_UOL (blue line). The broken magenta line represents a model simulation where the European fluxes from INV_ACOS inversion are replaced by INV_TCCON estimates.

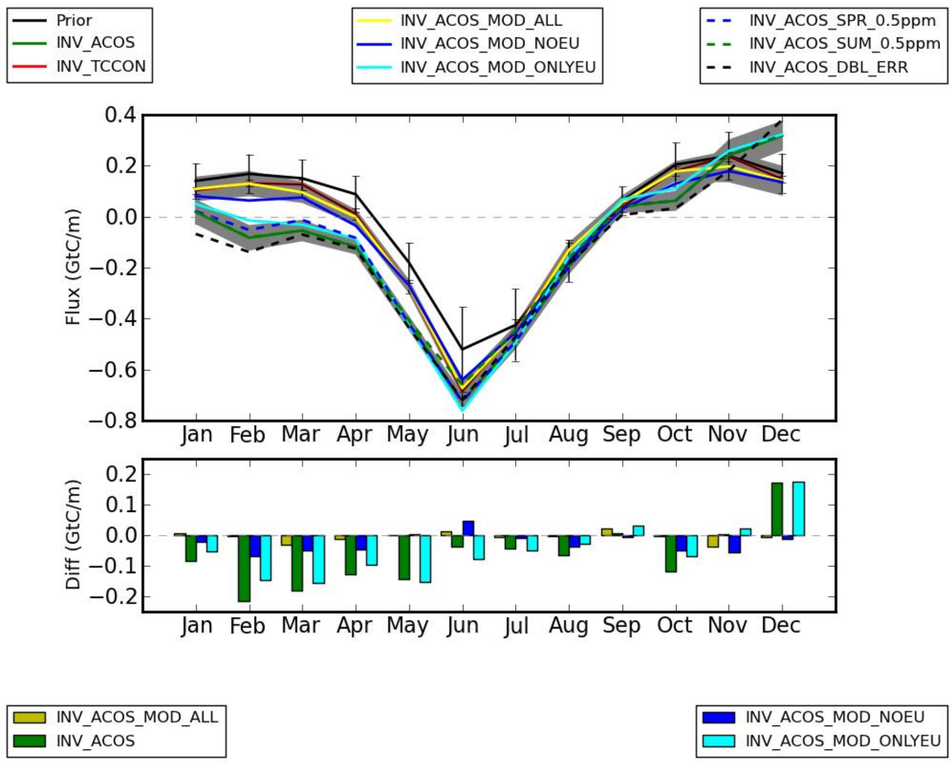

Figure 4: Monthly European biospheric flux estimates (GtC/m) from two groups of sensitivity experiments (top panel, Table 1). Black, green and red solid lines denote the a priori and the INV_ACOS and INV_TCCON inversions, respectively. Differences between INV_TCCON inversion and sensitivity inversions (bottom panel): 1) INV_ACOS_MOD_ALL (yellow), where all GOSAT retrievals are replaced by the model simulations forced by INV_TCCON a posteriori fluxes; 2) INV_ACOS (green), where original GOSAT ACOS retrievals are assimilated; 3) INV_ACOS_NOEU (blue) where all the GOSAT retrievals outside the European region are replaced by the INV_TCCON simulations; and 4) INV_ACOS_MOD_ONLYEU (cyan) where only GOSAT retrievals within the European region are replaced by the INV_TCCON simulations.

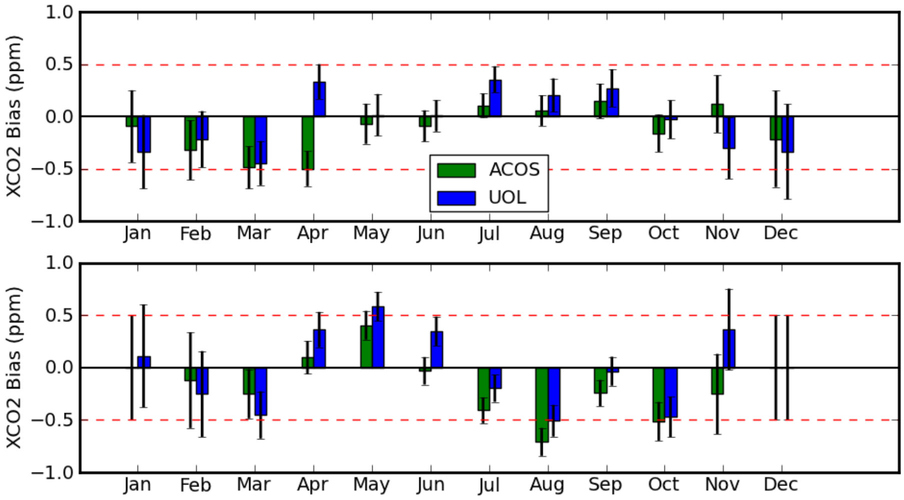

Figure 5: Estimates of monthly CO_2 biases (ppm) in GOSAT ACOS (green) and UOL (blue) X_{CO2} retrievals over (top) West (West of 20^{o}E) and (bottom) East (East of 20^{o}E) Europe. The black vertical lines represent the uncertainty.

Appendix A: Quasi-regional flux inversion.

To further study the contributions from X_{CO2} retrievals within and outwith Europe we have performed quasi-regional flux inversions to infer the European uptake of CO_2 in 2010, based on the same EnKF approach as the global flux inversions. In contrast to the global experiments (Table 1), for the quasi-regional inversions we assimilate observations only over Europe, and assign a small a priori flux uncertainty to any region outside Europe in order to minimize the influence of observations taken over Europe on other regions. Consequently, a posteriori flux estimates outside of Europe are close to their a priori values. We use the a posteriori fluxes from INV_TCCON as the a priori estimates for 12 sub-regions in Europe, and assume their uncertainty is two thirds of that we use for the global flux inversions. This is because the a posteriori estimates from INV_TCCON have already been refined by in situ data.

To investigate the influence of lateral boundary conditions on the quasi-regional flux inversions, we use two different sets of a posteriori estimates to define fluxes outside Europe: 1) INV_TCCON (INV_BD_TCCON) and 2) INV_ACOS (INV_BD_ACOS). Figure A1 shows that INV_BD_ACOS has a higher annual uptake of 1.58 GtC/a than INV_BD_TCCON with an uptake of 0.79 GtC/a (Table A1), with differences larger during the first half of 2010. The estimate for INV_BD_ACOS is similar to its global inversion counterpart INV_ACOS. Large differences between INV_BD_ACOS and INV_BD_TCCON highlight the importance of accurate lateral boundary conditions to a regional European inversion.

We use on-line bias correction schemes to reduce the adverse impacts from incorrect boundary conditions around Europe. Similar to Reuter et al. (2014), we estimate monthly observation biases across Europe using our quasi-regional flux inversion system. Here, we introduce a monthly bias to remove the systematic difference between model and GOSAT observations across the whole European region, and assume an associated a priori uncertainty of 100 pm (Reuter et al, 2014). This is different from our previous bias assumption of 0.5 ppm over East and West Europe for INV_ACOS_INS. Compared to INV_ACOS_INS, we also do not assimilate any in situ observations as additional constraints. Figure A1 shows that such a bias correction scheme (INV_BD_ACOS_BC) successfully reduces European uptake of CO_2 during 2010 to 0.96 GtC/a from 1.58 GtC/a for INV_BD_ACOS. Table A1 shows that after the applying bias correction scheme, INV_BD_ACOS_BC and INV_BD_TCCON_BC are consistent (0.94 GtC/a vs 0.96 GtC/a) despite different lateral boundary conditions provided by INV_ACOS and from INV_TCCON. But INV_BD_TCCON_BC (0.94 GtC/a) has 0.15 GtC/a more uptake than INV_BD_TCCON (0.79 GtC/a). We find a similar difference using UOL data (not shown), which infer an annual uptake of 0.71 GtC/a (0.56 GtC/a) with (without) the on-line bias correction.

We next examine the effectiveness of the inversion system that uses an on-line bias correction with large a priori uncertainty. Generally, large a priori uncertainty for biases will lead to the eventual loss of constraint by the observed mean CO_2 concentration across Europe. The weakened constraint can be seen by the enlarged a posteriori error (by 0.04 GtC/a) for INV_BD_TCCON_BC. In additional OSSEs (Table A2) we find that the loss of such a constraint can result in large systematic errors in estimated fluxes.

In these OSSEs, we assume the a priori estimates for 12 European sub-regions to be the same as the a priori used by INV_TCCON. Similar to INV_BD_TCCON, we set the fluxes outside European region to be the a posteriori estimates by INV_TCCON. We assimilate the INV_TCCON model ACOS X_{CO2} retrievals over Europe, to test the ability of the system to recover the "true" European flux (defined by INV_TCCON) from the assumed a priori that we define as the CASA model. Without the on-line bias correction, the quasi-regional inversion INV_REG_ENKF reproduces the truth for most months (Figure A2), and the associated annual uptake of 0.55 GtC/a compared to the true value of 0.58 GtC/a. If we also estimate monthly X_{CO2} bias with a large a priori uncertainty of 100 ppm (INV_REG_BC), the a posteriori European uptake is systematically underestimated for almost all months in 2010 (Figure A2). Consequently, the a posteriori annual uptake is about 0.38 GtC/a, which is 35% smaller than the truth (Table A2). Weakening the observation constraint also enlarges the a posteriori uncertainty from 0.22 GtC/a for INV_REG_ENKF to 0.27 for INV_REG_BC. But we find that increases in the estimated a posteriori uncertainty (by 0.05 GtC/a) are smaller than the increase in the systematic deviation from the true annual uptake (by 0.19 GtC/a).

More importantly, we find that the derived annual uptake is not linearly correlated to the assumed true fluxes. In experiment INV_REG_BC_SP (Table A2) we replace the true fluxes (defined by INV_TCCON) over the first 3 of 12 European sub-regions, which are at the south part of Europe (roughly south of 47^{o} N), with values from CASA model. As a result, the new true fluxes have an annual uptake of about 0.48 GtC/a across Europe, which is about 18% (0.1 GtC/a) lower than the original one defined by INV_TCCON for INV_REG_BC. We then re-generate model ACOS X_{CO2} data by running GEOS-Chem driven by the new hybrid true fluxes. However, after assimilating the new model X_{CO2} data, INV_REG_BC_SP infers an annual uptake of 0.37 GtC/a, which is almost the same as the posterior estimate (0.38 GtC/a) of INV_REG_BC, failing to reproduce the 18% decrease from the true value of 0.58 GtC/a assumed for INV_REG_BC to the 0.48 GtC/a assumed for INV_REG_BC_SP. In contrast, the quasi-inversion without on-line bias correction (INV_REG_ENKF_SP) well reproduces such decrease.

The bias correction across Europe can also increase the sensitivity to sub-regional biases. To illustrate this we added 1 ppm bias to the simulated observations during June to August of 2010 over south-west Europe between 35^{o}N to 42^{o}N and 15^{o}W to 20^{o}E (mostly over Spain and Italy). Without an on-line bias correction, adding the 1 ppm bias over the south-west strip leads to a small change (0.01GtC/a) in the annual uptake: a (slightly) reduced uptake in the first half of 2010 is largely compensated by a slightly enhanced uptake in the second half of 2010. Conversely, when we use an on-line bias correction with large prior errors (INV_REG_BC_1ppm), the 1 ppm positive bias increases the uptake by about 0.24 GtC in June, July and August. This implies that without the constraint from the mean concentration across the whole European region, the inversion system is free to interpret the higher concentrations over the small south-west strip as the signal of more uptakes over other larger part of Europe. As a result, the annual uptake changes from an underestimation of 35% by INV_REG_BC to an overestimation of 15% by INV_REG_BC_1ppm (0.65 GtC/a) (Table A2).

In summary, our quasi-regional inversion experiments highlight the sensitivity of regional flux inversions to the accurate description of the boundary conditions around the domain. Using an on-line bias correction can be helpful when the bias has been properly characterized. Over-correcting the bias can weaken the observation constraints, and possibly increase sensitivity to other small-scale unknown biases. We have also tested bias correction schemes using a different inversion algorithm (the Maximum A Posteriori (MAP) approach, Fraser et al., 2014), and found similar deficiencies when the a priori uncertainty of the regional observation bias is assumed to be very large. Our studies cannot prove or disprove Reuter et al. (2014), but it does highlight previously unrecognized limitation to the approach. The diversity of results reached under different assumptions associated with observation biases and emission spatial patterns highlight the importance for us as a community to investigate the interaction between observation and the inversion system for achieving consistent flux estimates in the future from assimilation of the up-coming observations from OCO-2 satellite as well as from the improved in situ networks.

Name	Description	Flux (GtC/a)	Uncertainty (GtC/a)
INV_BD_TCCON	Only ACOS data over Europe are assimilated to infer monthly fluxes over 12 European sub-regions. Fluxes outside EU are fixed to INV_TCCON inversion.	-0.79	0.18
INV_BD_TCCON_BC	The same as INV_BD_TCCON, but monthly bias with an assumed prior uncertainty of 100 ppm are included as additional parameters to be estimated.	-0.94	0.22
INV_BD_ACOS	The same as INV_BD_TCCON, but external regional fluxes are fixed to INV_ACOS.	-1.58	0.18
INV_BD_ACOS_BC	The same as INV_BD_ACOS, but estimates for monthly observation bias included.	-0.96	0.22

Table A1: The same as Table 1 but for quasi-regional inversions where only ACOS X_{CO2} are assimilated.

Name	Description	Flux (GtC/a)	Uncertainty (GtC/a)
INV_REG_ENKF	Synthetic ACOS data over Europe are assimilated to infer monthly fluxes over 12 European sub-regions, which prior estimates are assumed to be same as INV_ACOS (i.e., CASA model). Here we assume the true fluxes be a posteriori of INV_TCCON inversion.	-0.55	0.22
INV_REG_BC	The same as INV_REG_ENKF, but estimates for monthly bias are included as additional parameters.	-0.38	0.25
INV_REG_ENKF_1ppm	The same as INV_REG_ENKF, but 1ppm bias is added to the synthetic observations over a strip at south-west Europe for three months from June to August in 2010.	-0.54	0.22
INV_REG_BC_1ppm	The same as INV_REG_BC, 1ppm bias is added to the synthetic observations over a strip at south-west Europe for three months from June to August in 2010.	-0.65	0.25
INV_REG_ENKF_SP	The same as INV_REG_ENKF, but the 'true fluxes' over the first 3 of the 12 European sub-regions are replaced by CASA model values.	-0.47	0.22
INV_REG_BC_SP	The same as INV_REG_ENKF_SP, but with on-line bias correction with assumed prior uncertainty of 100 pm.	-0.37	0.25

Table A2: The same as Table A1 but for Observation System Simulation Experiments, where we assimilate synthetic ACOS X_{CO2} from model simulations forced by the assumed the 'True fluxes'.

640

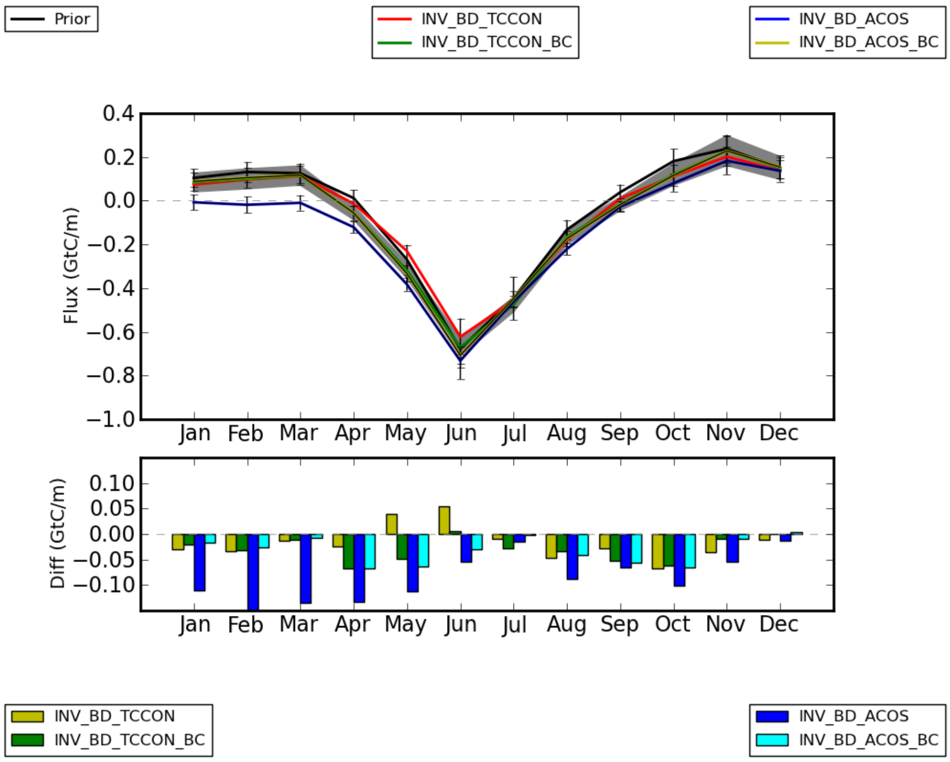

Figure A1: As Figure 4, but for the comparisons between the quasi-regional inversions. All the inversion experiments assimilate the same ACOS data set over Europe, with the a priori for 12 European sub-regions taken from posterior estimates from INV_TCCON. Fluxes outside Europe are fixed to the posterior estimates of INV_TCCON (INV_BD_TCCON and INV_BD_TCCON_BC) or to the estimates of INV_ACOS (INV_BD_ACOS and INV_BD_ACOS_BC). INV_BD_TCCON_BC and INV_BD_ACOS_BC also estimate the monthly bias across Europe as an additional parameter with an assumed a priori uncertainty of 100 pm estimated from ACOS data.

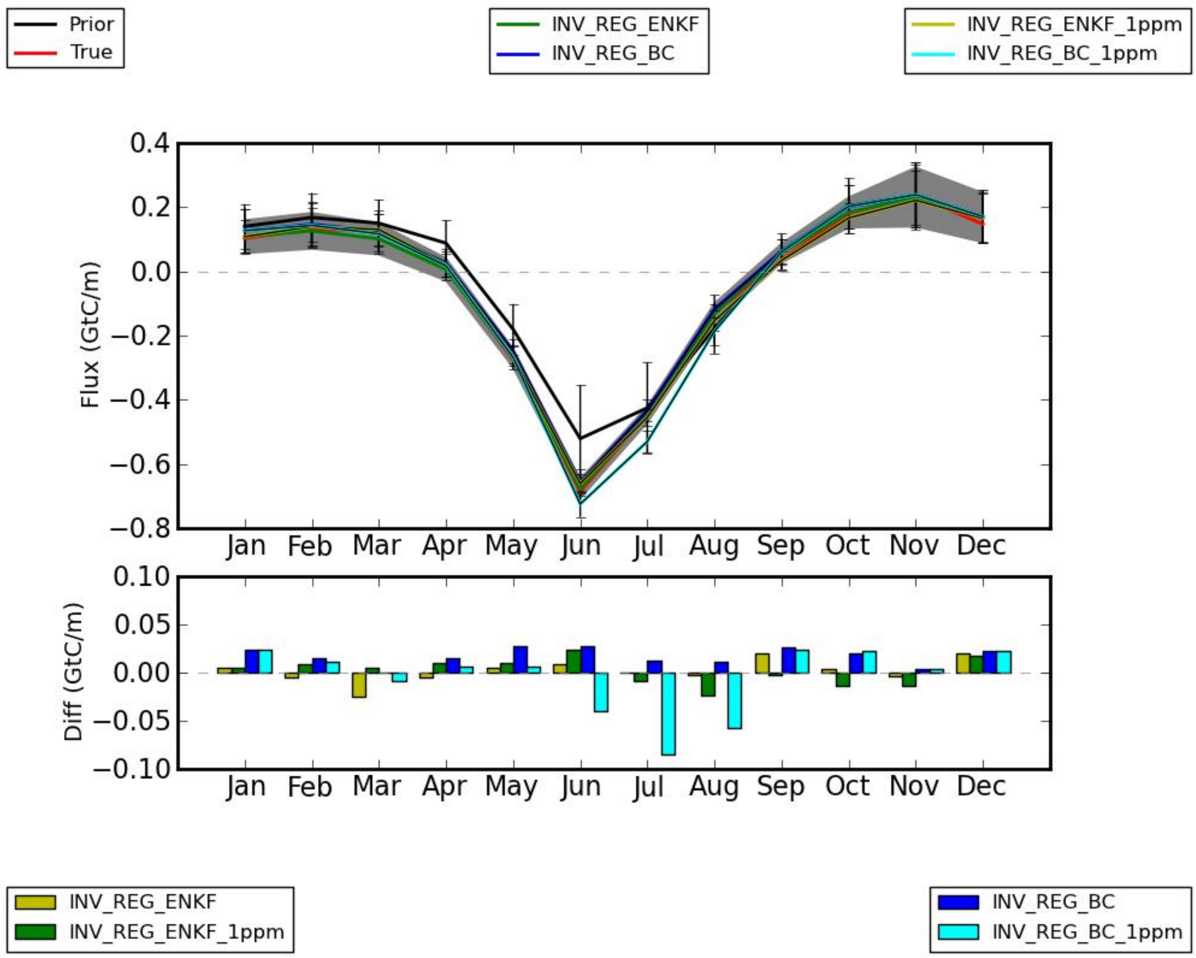

Figure A2: As Figure 4, but for comparisons of the quasi-regional inversions for assimilation of synthetic ACOS retrievals against 'True' fluxes (INV_TCCON). All the quasi-regional inversions have assumed the same a priori fluxes. But INV_REG_BC and INV_REG_BC_1ppm also include the monthly observation bias across Europe, with a prior uncertainty of 100 pm, as additional parameters to be estimated from the synthetic observations. In INV_REG_ENKF_1ppm and INV_REG_BC_1ppm, 1ppm observation bias is added to the (synthetic) observations over a small south-west strip of Europe during the summer of 2010.