Review of "Elevated uptake of CO_2 over Europe inferred from GOSAT XCO_2 retrievals: a real phenomenon or an artefact of the analysis?" by L. Feng et al

Overall impression

Over the past several years, CO_2 source-sink inversions ingesting satellite XCO_2 data from GOSAT (Basu et al., 2013; Chevallier et al., 2014; Deng et al., 2014; Reuter et al., 2014), SCIAMACHY (Reuter et al., 2014) and TES (Nassar et al., 2011) have all estimated a terrestrial CO_2 sink over Europe significantly in excess of what is estimated from inversions using only in situ (primarily surface layer) data. Since this satellite-derived sink is also much larger than what typical biosphere models estimate, there is an ongoing debate about whether the satellite-derived sink is "real" or an artefact of either remote sensing CO_2 data, or of the way inversions are set up to ingest them. If it is an artefact, it is likely to be a rather subtle one; Reuter et al. (2014) found the sink to be remarkably resistant to a host of sensitivity tests designed to test several likely artefacts in an inversion.

In this manuscript, the authors have entered the fray by running their own sensitivity tests – first by swapping out GOSAT XCO_2 over different geographical areas with simulated XCO_2 from an inversion based on in situ data, then by estimating a bias in GOSAT XCO_2 region by region – in order to settle the real versus artefact debate. I have reviewed the old manuscript, the comments by the previous three referees and responses from the authors, as well as the revised manuscript. In my opinion, the revised manuscript is certainly improved compared to the old one. In their comments, the authors have clarified that their aim is not to refute Reuter et al. (2014), but to provide an alternate hypothesis as to the origin of the enhanced European sink. They have also presented a few necessary additional experiments, such as what happens in a joint in situ and GOSAT assimilation when bias parameters are not optimized.

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 XCO_2 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 XCO_2 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 XCO_2 inversions", or something along those lines. I like those sensitivity tests, and I think they should be in the published literature for reference. 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. But then why do they say in the abstract that they can explain only 0.18 GtC/a from XCO_2 biases, when the enhancement they need to explain is 0.62 GtC/a?" 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 XCO_2. It is about demonstrating that certain spatiotemporal patterns of XCO_2 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.

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.

- **Line 199** INV_ACOS gets a European sink stronger by 0.82 GtC/a than INV_TCCON. The authors substitute ACOS XCO_2 outside Europe with the posterior CO_2 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 XCO_2 inside Europe, while the remaining 0.52 GtC/a strengthening is due to biases outside Europe[1].

  This logic behind this partitioning of the enhancement to XCO_2 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).

- **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 XCO_2 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 XCO_2. Same comment applies to lines 212-218; just because adding 0.5 ppm CO_2 has a certain effect does not mean that for the real data the same effect is the *result* of a 0.5 ppm bias.

- **Line 225** The way the authors have set up the biases in XCO_2 to be optimized is unphysical. Any possible bias in GOSAT XCO_2 is highly unlikely to vary with TRANSCOM regions. Rather, it may vary with geophysical 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 CO_2 data that can distinguish between the two.

  The number of bias parameters and their chosen prior errors are also troubling, to say the least. The 144 independent bias parameters/year make it very easy for the inversion to trade TRANSCOM-scale flux differences for TRANSCOM-scale XCO_2 biases, making both bias and flux estimates inaccurate. The choice of 1σ error of 0.5 ppm for the bias parameters is also baffling. The entire east-west gradient in XCO_2 across

[1] If I understand correctly, "bias" in XCO_2 here means difference from INV_TCCON posterior fields and not difference from some absolute truth.

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 XCO_2 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 XCO_2 which did not agree with INV_TCCON. Why even bother assimilating GOSAT XCO_2, in that case?

- **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 XCO_2 within Europe, even if the average GOSAT XCO_2 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 CO_2 field.

- **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 XCO_2 from INV_ACOS_INS/0.01 should match GOSAT XCO_2 everywhere, as well as in situ CO_2. This means that GOSAT XCO_2 – 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 XCO_2. 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 XCO_2 – 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 XCO_2 space, of 0.63 GtC/a flux from Europe.

  To what extent the two inversions match GOSAT XCO_2 is, of course, defined by the error settings on their in situ vs GOSAT observations. The authors have already presented their errors for GOSAT XCO_2 and TCCON XCO_2, 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, ...).

- **Line 267** The authors seem to have performed an inversion (not in their Table 1) where they've corrected XCO_2 over Europe by their estimated bias correction and left the other XCO_2 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.

- **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 XCO_2. I can think of another possibility. What if the CO_2 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?

  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.

Minor point-by-point comments

- **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."
- **Line 72** Please replace "emission" by "source". In typical usage in literature, the former includes, e.g., fossil fuel combustion but not ecosystem respiration.
- **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 XCO_2 has more far field influence. Have the authors considered that?
- **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.
- **Line 125** Calculating the error in TCCON XCO_2 as the variation of XCO_2 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 XCO_2, the forward model probably captures some of the variation in XCO_2 between 9:00 and 15:00 LT, which should therefore not be part of the error assigned to the observation.
- **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 CO_2 over the tropical Pacific, but they haven't made it clear why that's important in this context.
- **Line 145** I do not see how INV_ACOS and INV_UOL fitting GOSAT XCO_2 better than INV_TCCON is any indicator of performance. The former two, after all, ingest GOSAT XCO_2, so they *should* fit GOSAT XCO_2 better! At best, this is proof that the machinery of the inversion system works.

- **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.

- **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 XCO_2 soundings outside Europe result in an unrealistically high inflow of CO_2 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 CO_2 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.

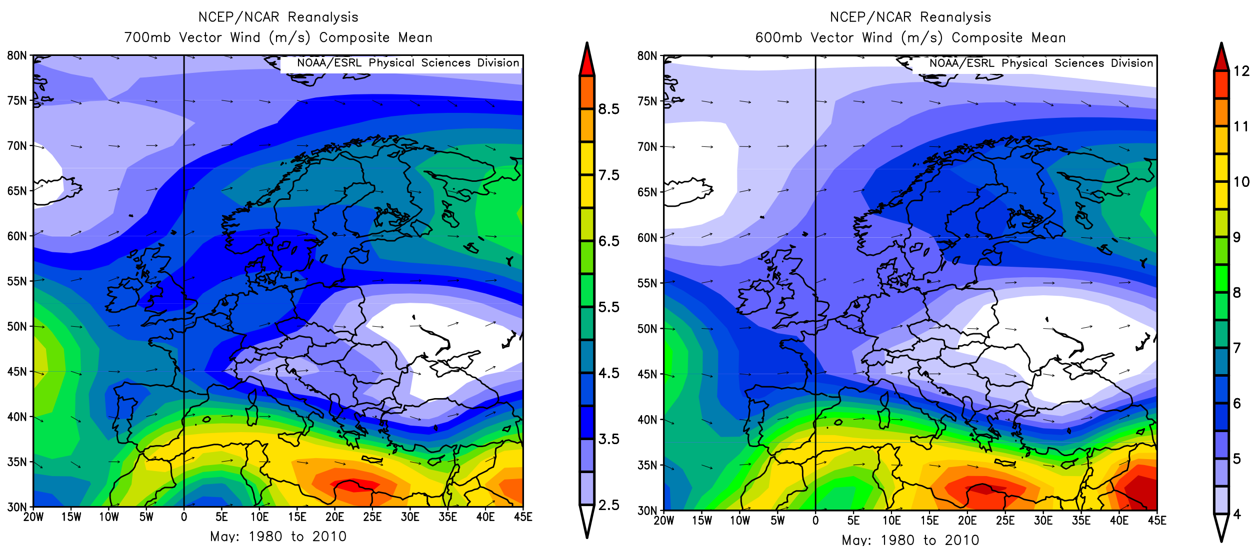

Figure C1: Climatological mean (1980 to 2010) wind vectors at 700 mb and 600 mb over Europe in May. Data from `http://www.esrl.noaa.gov/psd/cgi-bin/data/composites/printpage.pl`

- **Line 194** The fact that INV_ACOS_MOD_ALL reproduces INV_TCCON reflects that the machinery of the inversion system ingesting XCO_2 data works. Not sure how it reflects the seasonal variation in GOSAT coverage.

- **Line 199** Why did the authors not present the converse experiment of INV_ACOS_MOD_NOEU, namely, INV_TCCON posterior XCO_2 inside Europe, actual GOSAT XCO_2 outside Europe?

- **Line 225** In the setup with bias parameter optimization, does a positive bias mean that GOSAT XCO_2 is too low, or too high? As in, what is supposed to be zero if all the data are fit perfectly, $XCO_2^{model} - XCO_2^{GOSAT} - \text{bias}$, or $XCO_2^{model} - XCO_2^{GOSAT} + \text{bias}$? Please write down the explicit equation(s) relating the bias-corrected and uncorrected XCO_2.

- **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?

- **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.

- **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 XCO_2.

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 estimated from GOSAT retrievals of total column CO_2, Atmospheric Chemistry and Physics, 13, 8695–8717, doi:10.5194/acpd-13-4535-2013, URL http://www.atmos-chem-phys-discuss.net/13/4535/2013/, 2013.

Chevallier, F., Palmer, P. I., Feng, L., Boesch, H., O'Dell, C. W., and Bousquet, P.: Toward robust and consistent regional CO2 flux estimates from in situ and spaceborne measurements of atmospheric CO2, Geophysical Research Letters, 41, 1065–1070, doi:10.1002/2013GL058772, URL http://dx.doi.org/10.1002/2013GL058772, 2014.

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, Atmospheric Chemistry and Physics, 14, 3703–3727, doi:10.5194/acp-14-3703-2014, URL http://www.atmos-chem-phys.net/14/3703/2014/acp-14-3703-2014.html, 2014.

Nassar, R., Jones, D. B. A., Kulawik, S. S., Worden, J. R., Bowman, K. W., Andres, R. J., Suntharalingam, P., Chen, J. M., Brenninkmeijer, C. A. M., Schuck, T. J., Conway, T. J., and Worthy, D. E.: Inverse modeling of CO_2 sources and sinks using satellite observations of CO_2 from TES and surface flask measurements, Atmospheric Chemistry and Physics, 11, 6029–6047, doi:10.5194/acp-11-6029-2011, URL http://www.atmos-chem-phys.net/11/6029/2011/, 2011.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO2 and implications for carbon cycle research, J. Geophys. Res., 109, D02 301, doi:10.1029/2003JD003968, URL http://dx.doi.org/10.1029/2003JD003968, 2004.

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, Atmospheric Chemistry and Physics, 14, 13 739–13 753, doi:10.5194/acp-14-13739-2014, URL http://www.atmos-chem-phys.net/14/13739/2014/, 2014.

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, Geophysical Research Letters, pp. n/a—-n/a, doi:10.1002/2013GL059174, URL http://dx.doi.org/10.1002/2013GL059174, 2014.