Point-by-Point Response to Referee Comments, Including Relevant Changes Made in the Manuscript, and Marked-Up Manuscript and Supplement

J. S. Wang et al.

Referees' comments below are in *italics*, and our responses are in regular text style.

Anonymous Referee #1

This paper, entitled "A global synthesis inversion analysis of recent variability in CO2 fluxes using GOSAT and in situ observations" by Wang et al., combines satellite and ground-based observations of CO2 concentrations to infer sources/sinks of carbon over the world.
One of the unique aspects of this work is the adoption of a batch Bayesian synthesis inversion approach that is at higher spatial (perhaps; see comments below) and temporal resolution relative to conventional global inversions that allow full-rank posterior error correlations to be evaluated. However, I have major questions about the results, and I found the conclusion about GOSAT observations to be weak. Therefore, I recommend major revisions prior for consideration for publication in ACP.
We thank the reviewer for the insightful comments. We have made major revisions as detailed below.

MAJOR COMMENTS I. Utility of satellite observations The authors carried out extensive comparisons between inversion results using in-situ data versus those from using the GOSAT satellite data, as well as both data. The results often differed dramatically, and the comparisons against independent aircraft data highlighted potential biases in the GOSAT data. I am left wondering whether it is worth assimilating the GOSAT data at all, given the fact that the inversion-derived spatial and temporal patterns of carbon fluxes could be highly erroneous, with significant biases that are not reflected in the posterior errors (see point II. below). After highlighting several issues with GOSAT in the Discussion and Conclusions section, the final paragraph was anticlimactic and weak. Can the authors provide a broader guidance on how the GOSAT data can be properly used (if at all)? Is there an important take-home message that this study can provide to the reader? A clear, concise message is lacking in this lengthy paper.

We now highlight the following take-home message: Our inversions demonstrate that passive satellites such as GOSAT can in principle reduce flux uncertainties relative to in situ networks because of their greater coverage, and that in situ and satellite observations complement each other. However, accurate estimates are currently hampered, especially on subcontinental/sub-ocean basin scales, by retrieval biases and remaining coverage gaps. With the current bias-corrected GOSAT data, interannual variations seem to be captured over certain regions, though still with some bias. To remedy the situation, bias corrections would need to be improved, especially over regions that are not currently sampled by TCCON and are challenging for forward model simulations, e.g. Africa. In addition, future expansions of in situ networks and coverage by potential active and passive satellite sensors would be helpful in filling gaps in joint in situ-satellite inversions.

In addition, we have strengthened the Discussion and Conclusions section through re-organization, e.g. by moving the material in the anticlimactic final paragraph to the end of the first paragraph of the section.

II. Uncertainties (prior and posterior errors) Several issues arose as I considered how

the prior and posterior errors were dealt with in this paper: 1) More quantitative information regarding the prior "observation" uncertainties at different sites based on the "model-data mismatch" would be helpfulâˇA¨Te.g., Table or a map of the prior errors at different sites.
We have now created a table (Table S1 in the supplement) containing the model-data mismatch errors for the in situ sites (as well as basic site information such as latitude, longitude, etc.).

2) The handling of transport errors requires more discussion.
The authors adopted a single transport model (instead of an ensemble of transport models as some other studies do), PCTM, and assumed that the "model-data mismatch" accounts for the transport error. How did PCTM perform as part of previous TRANSCOM experiments, and other tracer experiments (e.g., SF6)? Its interhemispheric mixing? How about its vertical mixing and rectifier strength?
We now elaborate on PCTM's transport performance, including in the context of TransCom experiments, in Section 2.3 (Atmospheric transport model and model sampling). The new text is as follows:
Evaluation of PCTM over the years has shown it to a reliable tool for carbon cycle studies. For example, Kawa et al. (2004) showed that the SF_6 distribution from PCTM was consistent with that of observations and of the models in TransCom 2, suggesting that the interhemispheric and vertical transport were reasonable. PCTM performed well in boundary layer turbulent mixing compared to most of the other models in a TransCom investigation of the CO_2 diurnal cycle (Law et al., 2008). The TransCom-CH_4 intercomparison (Patra et al., 2011) showed that a more recent version of PCTM performed very well relative to observations in its interhemispheric gradients of SF_6, CH_3CCl_3, and CH_4 and interhemispheric exchange time, and follow-on studies (Saito et al., 2013; Belikov et al., 2013) demonstrated through evaluation against observed CH_4 and ^{222}Rn that the convective vertical mixing in PCTM was satisfactory overall.

Where does PCTM fit within the well-known "diver-down" figure in Schimel et al. [2015] (www.pnas.org/cgi/doi/10.1073/pnas.1407302112; its Fig. 3)?

We have now calculated the north-south land flux partitioning for our various inversion cases and display them in a diver-down plot (shown below, and now included in the Supplement for our paper). We find that our results for the in situ inversions and especially the GOSAT inversions lie outside of the 1-sigma range consistent with the GCP global land flux estimate, with rather large positive values (carbon source) for the south+tropics. An exception is the tight-priors case for the in situ inversion, which lies on a 1-sigma boundary. One reason for the weak overall land sink in our results is the large ocean sink compared to GCP for all cases, which is a consequence of our loose prior ocean fluxes and necessitates a small land sink for global mass balance. In addition, the GOSAT inversion total budget over relatively short 12-month periods may deviate substantially from a budget based on in situ measurements as in GCP (for the reasons proposed at the beginning of Section 3.2 of our manuscript). We also find that the north-south partitioning is somewhat sensitive to the time period considered, which could be due to the short 12-month periods here. For example, the Sep 2009-Aug 2010 period gives a smaller source in the south than the Jun 2009-May 2010 period for the GOSAT inversion, by ~0.7 Pg C, with a correspondingly smaller sink in the north.

We now refer to the diver-down plot (Fig. S2) in Section 3.2 of the manuscript in discussing the shift in the sink from the south + tropics to the north in the GOSAT inversion relative to the in situ inversion, as well as in several places in Section 3.3 in discussing the land-ocean flux partitioning and sensitivity to prior uncertainties.

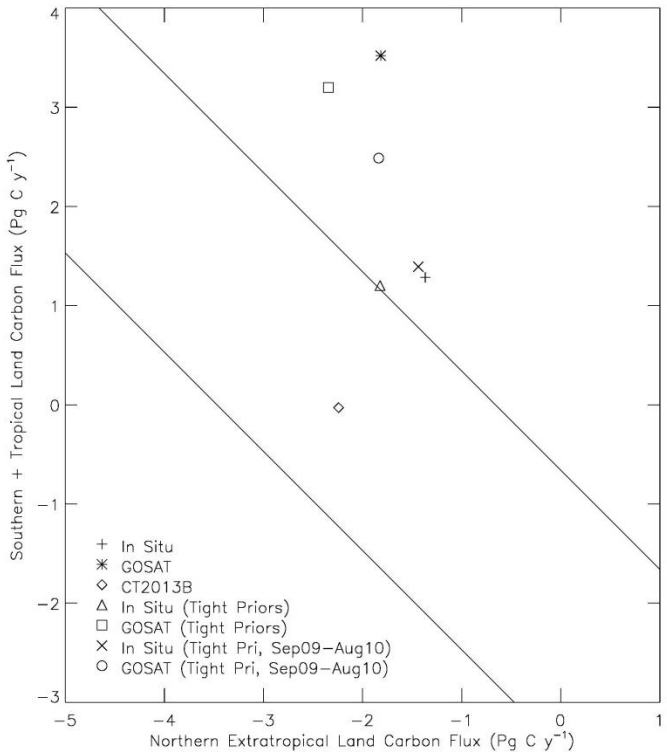


Figure S2. Posterior north-south land flux partitioning after Schimel et al. (2015). The diagonal lines are based on the global land carbon exchange (= land-use change emissions – land sink) estimated by GCP (2015) for the years relevant to the present analysis, i.e. 2009 and 2010, ± 1σ. Fluxes are for June 2009-May 2010 except where specified in the legend (for September 2009-August 2010). CT2013B refers to the CarbonTracker data assimilation system.

3) What is the logic behind neglecting a priori spatial and temporal error covariances? The stated reasoning is "The neglect of a priori spatial error correlations is justified by the size of our flux optimization regions, with dimensions on the order of one thousand to several thousand km, likely greater than the error correlation lengths for our 2_ x 2.5_ grid-level fluxes. For example, Chevallier et al. (2012) estimated a correlation e-folding length of ~500 km for a grid size close to ours of 300 km x 300 km based on comparison of a terrestrial ecosystem model with global flux tower data." If the inversion for flux adjustments is carried out at scales of ~1000 x 1000 km, following the TRANSCOM regions, this is imposing an artificial error covariance lengthscale of ~1000 km (significantly larger than the ~300 km lengthscale), thereby resulting in aggregation errors. Also, given this coarse lengthscale, can the authors make the claim that the adopted inversion is really high spatial resolution? If 2_ x 2.5_ grid is considered high resolution, this is not much different from CarbonTracker's spatial gridding.

We acknowledge that the size of our flux optimization regions, though smaller than that of previous batch inversions, does still impose a correlation that can result in aggregation error. Note that the proper correlation length scale is actually ~500 km, not ~300 km, and the number refers to e-folding, so the correlation is not negligible beyond 500 km. Regarding our claim about the "relatively high spatiotemporal resolution" of our inversion in line 110, we were specifically comparing the *flux optimization* resolution to that of previous *batch* inversions, e.g. TransCom 3 and Butler et al. (2010). We now make that more clear in the paper. We similarly clarify the statement about the "high spatiotemporal resolution of our inversion relative to other

Bayesian synthesis inversions" in lines 150-151. Actually, the spatiotemporal resolution of our flux optimization (108 regions, 8-day) is comparable to or even higher than some Kalman filter/smoother systems in use, e.g. CarbonTracker (156 regions, weekly), Feng et al. (2009; 144 regions, 8-day), and Takagi et al. (2011; 64 regions, monthly). As for the transport grid resolution, we did mention in lines 212-214 that "the relatively high-resolution transport model used here (Sect. 2.3) captures much of the variability in the observations beyond background levels." We have now changed "relatively" to "reasonably" since how the transport resolution compares to other models is irrelevant in this statement.

4) Are the prior uncertainties of proper magnitude? While the authors enlarged the observation uncertainties by a factor of 2 to lower the normalized Chi-squared value closer to 1, Table 2 shows that the in situ only inversion results in a Chi-squared of 4.0, which suggests that the prior uncertainties may still be underestimated. Should the prior uncertainties be inflated even more?
We wanted to strike a balance between achieving a chi-squared value reasonably close to 1 and not inflating the uncertainties derived from our in situ data uncertainty formulation too much. In addition, in a test in which we further inflated the uncertainties to 3 times the original values, the posterior fluxes were not very different from those for the reported inversion at the scale of TransCom regions, so that our results appear to be robust. Finally, the use of the chi-squared criterion to scale the prior and/or observation uncertainties is inherently a rough approach, as pointed out by Rayner et al. (1999, *Tellus*). We now add a brief statement in Section 2.2 of the manuscript mentioning the overall robustness of the results with respect to observation uncertainties.

5) Can we place much faith in the posterior errors for calculated sources/sinks that are derived through the inversion? There are several cases in the paper where different inversions yield different carbon fluxes whose values do not overlap even if allowing for the posterior error bars. Since there is only one true carbon cycle, the posterior errors should allow for the "true fluxes" to be encompassed within the error bars even under different inversion setups. How much trust (if any), therefore, should we place on the posterior errors? I realize that other global inversion studies often face the same situation of underestimating posterior errors, but this would be an opportunity for the authors to comment on the posterior errors.

The Bayesian, least-squares inversion framework adopted here, as in other CO_2 studies, assumes Gaussian error distributions with no bias (observation, transport, prior, etc.), so the posterior uncertainties do not account for possible biases. That's why we discussed in much depth the possible biases in the results, using independent observations for evaluation. We now explicitly state this limitation of the posterior uncertainties after Eq. 3 for the posterior error covariances in Section 2.4. Given that limitation, one wouldn't necessarily expect posterior error bars to overlap for different inversions, e.g. in situ vs. GOSAT, which may have different regional biases.

Note that we mentioned overlaps (or lacks thereof) mostly at the 1-sigma uncertainty level, as that is what the error bars in our figures represent. But two rather different flux estimates could still be considered to encompass the "truth" if they overlap at a 2-sigma level, for example. Also, in a comparison of monthly fluxes from our in situ inversion and CT2013B, we highlighted the similarity, not difference, of the two sets of fluxes: "with overlapping 2σ ranges at all times except in the Northern Oceans region" (lines 491-492). In another part of the paper, we mentioned large differences that even exceed 3-sigma (for Europe), but that refers to comparison of two different years--2010 and 2009 summer posterior fluxes.

III. European sink The large European sink of -1.5 PgC/yr is difficult to reconcile with ground-based information. While the large sink could be caused by biases in GOSAT data, as pointed out by the authors, I am puzzled by the fact that the summertime drawdown is so large in the prior fluxes as well, reaching fluxes of -5 PgC/yr and comparable with the summertime drawn in boreal and temperate North America. Is there really enough biomass/vegetation to sustain a summertime drawdown of -5 PgC/yr in Europe? Also, I suggest the authors consider and cite a recent paper focusing on the European sink [Reuter et al., 2017, BAMS; http://dx.doi.org/10.1175/BAMS-D-15-00310.1].

Our GOSAT inversion results are not unique, as the average GOSAT-only posterior annual flux for Europe over the different models in the intercomparison of Houweling et al. (2015) was around -1.5 Pg C/y as well. As for our prior fluxes, comparing them with the mean seasonal cycle over the early 2000s for a set of in situ inversions reported by Peylin et al. (2013) shows that our prior for Europe falls squarely within the range of inversions, for example peaking at ~-0.5 Pg C/month in the summer (equivalent to ~-6 Pg C/y). (There is agreement for North America and North Asia as well.) Also, the TransCom 3 Europe region includes western Russia, so it is actually of comparable size to Boreal or Temperate N. America and contains much forested land.

We now cite Reuter et al. (2017) in our discussion of the shift in the global sink in the GOSAT inversion relative to the in situ inversion, noting that they highlight a similar discrepancy between satellite-based and ground-based estimates of European CO2 uptake and cite retrieval and sampling biases as possible sources of error in the former (while also noting sampling issues with in situ networks for the region).

OTHER COMMENTS Length of text: This is a long paper, with the main text reaching over 40 pages (double-spaced). The Results section as a whole felt unnecessarily long. Can the authors seek to condense?

We now omit some details/condense the text especially in the Results section, for example in the evaluation of the model against Amazonica profiles (omitting details on numbers of cases in which one inversion is better than another in lines 404-412), and in the evaluation against HIPPO (lines 566-597). We have also moved Figures 13 and 14 (comparisons with HIPPO and surface observations for tight priors case) to the Supplement (as Figures S3 and S4), saving some space in the main document. Although the remaining text is still lengthy, we have put much effort into making the wording as concise as possible and the overall paper is comparable in length to other global CO2 inversion papers that assimilate satellite and in situ data and use multiple evaluation data sets.

Spatiotemporal resolution: From the start of the paper the fact that the inversion is at "high resolution" is mentioned numerous times. This is supposedly a strength of the inversion technique adopted in this study. However, it is not until deep in the Methodology section that the reader finds out exactly what resolution is adopted in both space and time. Can the authors specify exactly what the resolution is early on (even in the Abstract)? Also, can the spatial gridding really be considered to be high resolution (see comment above in comparison to CarbonTracker)?

As described in our response to the earlier comment, we have revised statements in the Introduction and in the beginning of Section 2 to clarify that "high resolution" refers to the flux optimization and is relative to previous batch inversions. Also in the beginning of Section 2, we provide some background on previous batch inversions. We thus believe there is a sufficient

level of detail early on to communicate the key point. Also as described earlier, we have now modified the statement about the relatively high grid resolution.

Lines 462-464: It appears that the in-situ only inversion also produces winter-time uptake in Boreal Asia (Fig. 6). This is not just a feature in GOSAT-based inversion. Why?
It's true that the in situ inversion exhibits a flux with a 1-sigma range entirely below zero in Boreal Asia in December. Given that the in situ inversion is generally noisier than the GOSAT inversion, as we pointed out in line 457, we considered the GOSAT feature more notable and focused on that.

Figs 8 and 12: The different colored bars are hard to distinguish. I suggest separating out the Global flux numbers as a separate panel and zooming in the figure to a smaller range for the regional results.
OK, we have divided each figure into 2 panels as suggested.

Anonymous Referee #2

General comments
This paper describes an impact of satellite observation data on carbon cycle inversion by using multiple settings (observation data sets and prior flux uncertainties) of high-resolution batch Bayesian inversion. The new results are well considered (ex. inversion bias can vary with data coverage). I consider this article should be acceptable after some minor revisions for publications for ACP. One important issue is that the number of observation sites (87 sites (only 10 continuous sites for in situ inversion) is considered insufficient to constrain 108 regional CO2 flux. The inversion results (chisquared value, dipole behavior, mismatch against independent observation, etc.) show this issue. One option to avoid this issue is to add observation sites (JR station data and amazonica aircraft data) for in situ inversion. The other issue is inadequate description of satellite retrieval bias. I consider the difference between in situ inversion and GOSAT inversion comes from not only satellite sampling bias but also satellite retrieval bias. The authors should discuss retrieval bias of satellite from validation of multiple inversion results and show some choices. Because modification of satellite biases (sampling and retrieval) is significantly important to the future use of satellite observation data in carbon cycle analysis.

We thank the reviewer for the insightful comments and the positive overall evaluation. We agree that using a larger number of sites would allow for better constraints on fluxes and reduce some of the posterior flux error correlations seen in the in situ inversion. Note that the number of sites, 87, is similar to that of CarbonTracker 2013B, as we mentioned in the paper, and actually larger than most of the models in the Houweling et al. (2015) inversion intercomparison, many of which had even higher flux optimization resolution than our inversion. The reviewer's suggestion of using JR-STATION and Amazonica data in the inversion is a good one (though the latter observations were made during only a part of our analysis period and at a lower frequency--about once every 2 weeks--than the 8-day resolution of our inversion). In fact, we pointed out in our Discussion and Conclusions that those data sets, and others, could also be incorporated in the inversion. We had instead reserved them for use as independent evaluation data. For example, in the discussion of the 2010-2009 boreal summer flux differences, we chose to focus on the constraints provided by GOSAT (the in situ inversion is not shown in Figs. 16 and 17 and Table 3), and used JR-STATION data to evaluate the GOSAT

inversion over Siberia. Note that some regions of the world, such as tropical Africa, are not well sampled by any long-term in situ networks. A final comment is that we delineated the flux optimization regions based not only on the distribution of in situ sites but also on the higher density of the GOSAT observations, with the expectation that the latter would provide greater reductions in flux uncertainties. We have now added the phrase, "which allows us to take advantage of the relatively high density of the GOSAT observations", in Section 2.4 after describing our large number of regions compared to previous batch inversions.

As for retrieval bias, we now cite that in more places than in our original manuscript (details are given below). Note that effects of retrieval bias and sampling bias are often intertwined in inversion results. Since there is a lack of evaluation data in certain important regions, e.g. Africa, we highlight this as a need for improved carbon cycle studies in the future in the Discussion and Conclusions.

Specific comments

Line 38: The authors should also mention the influence of satellite observation errors on this sentence.

In this part of the abstract, we are discussing how the coverage of the in situ and GOSAT data sets affects their ability to constrain fluxes, as reflected in posterior flux uncertainties and correlations. Satellite retrieval errors do not affect the inversion-generated error covariance estimates, since the Bayesian, least-squares method used assumes Gaussian error distributions with no bias. Thus, the retrieval errors are not relevant in the statement about the spatial scales at which the data sets can resolve fluxes. However, we do discuss the effects of satellite retrieval biases soon after that in lines 42-44.

Line 54: At 2010, a relatively high temperatures around eastern North America event (all year round) occurred. This event might have affected greater uptake over the region.

High temperatures could indeed have contributed to greater uptake, especially at higher latitudes, where insufficient warmth could be more of a limiting factor for NEP than insufficient moisture during late spring-early summer. We looked at plots of the 2010-2009 flux difference in May-June for the inversions, and did see some areas of increased uptake in 2010 at high latitudes (including in the east), although the geographic pattern was not entirely consistent across the prior, in situ posterior, and GOSAT posterior. We now mention in Section 3.4 the possible contribution of higher temperatures to increased uptake especially at higher latitudes.

Line 68: The authors should refer to the high precision feature of observation data in this sentence.

OK, we have modified the sentence after that one to include that feature: "However, the accuracy of top-down methods is limited by incomplete data coverage (especially for highly precise but sparse in situ observation networks),…." (We modified that sentence rather than the one in line 68, given that high precision is not a general feature of top-down estimation, applying only to in situ mole fraction observations and not to current satellite retrievals.)

Line 178: In the experimental settings, natural (biosphere and ocean) net sink is too small comparing with current knowledge. This means that it is necessary to redistribute a large amount of CO2 flux by inversion. Errors tend to occur due to transport model, observation data and inversion settings. The author should mention about it.

It's true that the prior fluxes should ideally be consistent with the best available knowledge on a global, annual basis, e.g. a sizable net land sink rather than close to neutral. Note however that the uncertainties we assigned to the prior fluxes are large relative to the possible prior underestimates of the net sinks (especially for land). It can be seen in Fig. 8 for example that

the 1-sigma range for the prior global land flux easily encompasses the CarbonTracker posterior estimate of -2.3 Pg C y^{-1} as well as our own posterior estimates using the in situ and GOSAT data sets. In fact, the GOSAT and in situ + GOSAT inversions infer a larger *source* than the prior, so that CO2 is actually redistributed in the opposite direction. And for the global oceans, the 1-sigma range for our prior encompasses the CT posterior. Interestingly, the prior is not strong enough of a constraint to prevent any of our three inversions' ocean fluxes from exceeding the lower bound of the prior range. We now point out in the manuscript that the prior uncertainties are large enough to accommodate possible biases in the prior fluxes.

Line 204: 87 sites seem to be insufficient comparing with 108 regions (numbers of continuous sites are only 10 and geographically unevenly distributed). The authors should use more observation sites to constrain these regional CO2 flux.
Please see our response to the general comments above.

Line 226: The minimum value (0.01ppm) of the standard deviation of the observations within a particular looks too small. The authors should clarify the reason.
We have now added the phrase "to avoid uncertainty values of 0" for explanation.

Line 307: The authors should show a thickness of the lowest model layer.
We now specify that the thickness of the lowest layer is ~100 m on average.

Line 308: The authors should show numbers of dimensions of matrix (especially GOSAT and in situ + GOSAT inversion).
OK, we now show the dimensions for the GOSAT and in situ + GOSAT inversions (at the end of Section 2.4).

Line 392: The authors should also mention the bias of satellite observation.
We were actually referring to bias when we stated that "the in situ and GOSAT data sets are not necessarily consistent with each other", and earlier in the same sentence, "Comparison of posterior mole fractions with the data set not used (Fig. 3b, c), on the other hand, gives mean differences not as close to 0 as in the comparison with the assimilated data". Elsewhere in the paper, we make it clear that the bias occurs in the satellite observations.

Line 398: The authors should discuss the reason why chi-square of GOSAT inversions is almost same.
The chi-squared for the GOSAT tight-priors case is slightly larger than that for the baseline inversion, 0.823 vs. 0.778 (the tight-priors case for the in situ inversion also has a larger chi-squared than the baseline), but the numbers both round to 0.8 for one decimal place, which is what we show in the Table. We now add a note in the discussion of the results for the tight-priors cases (in Section 3.3) about the difference in chi-squared (cost function value) being concealed by the rounding.

Line 422-425: The authors need multifaceted discussions. It is difficult for this inversion setting (PCTM horizontal resolution (about 200km) and 8 day mean flux) to reproduce CO2 concentrations near surface at regions where vegetation activities are active like tropical rainforests.
Our prior biospheric fluxes actually have 3-hourly temporal resolution and our transport model has hourly resolution, so the model can potentially reproduce some of the short-term variability in the aircraft observations. (We now add text to the transport model description section to highlight the hourly resolution.) But it is true that our inversion is designed to improve the fit to

observations at broad spatial and temporal scales and not at small scales. One factor that likely contributes to the posterior model-observation differences having greater variance than those for the prior at low altitudes is that the GOSAT data are sparse over the Amazon, so that there is little data averaging over the 8-day intervals and flux regions. Thus, random errors can have a substantial impact. We failed to mention this in the text, so we now add it. Also, our original statement about the GOSAT random errors not being correlated with those of the aircraft observations is obvious and actually irrelevant, since the aircraft observations, based on flask samples, have negligibly small measurement errors. We now omit that statement.

Line 434-437: The authors should show the usefulness of increasing the number of observation data.
The focus in this paper is on a comparison of inversions using satellite and in situ data. As such, we believe it is sufficient to show the usefulness of the larger number of observations provided by GOSAT without presenting multiple in situ inversions using different numbers of observations. Please also refer to our response to the General Comments above for a discussion of the number of sites used in the in situ inversion.

Line 448-453: The authors should show degrees of freedom for signal and noise for in situ + GOSAT inversion as previous paragraph.
OK, we have now added those numbers and some text commenting on them.

Line 537: The authors should unify expression (prior/in situ/GOSAT) for N. Pacific and N. Atlantic.
We intentionally left out the priors for those regions because the GOSAT inversion differs from only the in situ inversion and not the prior over northern oceans, as stated in the sentence before that line. We wanted to show only the relevant numbers here.

Line 572: The lack of ocean observations at southern high latitudes brings analysis results closer to a priori information. The authors should consider satellite retrieval bias.
Our original thinking was that the underestimate of the GOSAT inversion relative to HIPPO in the southern extratropics reflected a negative correlation of the posterior fluxes with excessively positive fluxes in the tropics, which is caused by insufficient observational constraints. But now we agree that retrieval bias likely plays a role in the region, since the magnitude of the discrepancies relative to HIPPO, as much as several ppm, is probably too large to be caused solely by sampling bias. Note that sampling bias still plays a role (via error correlations), since the posterior mole fractions are worse than the prior ones even where there are no GOSAT observations. We now add retrieval bias as a factor in the text here.

Line 577-579: 67ppm difference seems too large. The author should identify and remove the cause observation data from inversion.
We agree the difference is unreasonable but do not find specific reason to omit the comparison for that particular latitude bin. We suggested that both the in situ and GOSAT inversions may be *under-constrained* in that region and season (high-latitude North Pacific and Alaska in the fall), so removing data in the inversions would be unlikely to help. Also, we pointed out that the prior is substantially higher than HIPPO here as well, and suggested that transport error might be a common cause for all the discrepancies.

Line 586-588: In general, current transport models could not well reproduce tropopause. The authors should use only tropospheric HIPPO data in figure 10 (c,f) for

discussion.
Although many transport models have problems in simulating the tropopause, we expect PCTM to perform well, given that it has high vertical resolution (with more than 20 levels in the stratosphere) and the tropopause in the underlying GEOS-5/MERRA meteorological data assimilation system is considered to be accurate. Evidence is provided by the TransCom-CH_4 evaluation of models of Patra et al. (2011) and the ozone tropopause transport analysis of Wargan et al. (2015). (Wargan, K., S. Pawson, M. A. Olsen, J. C. Witte, A. R. Douglass, J. R. Ziemke, S. E. Strahan, and J. E. Nielsen (2015), The global structure of upper troposphere-lower stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data, J. Geophys. Res. Atmos., 120, 2013–2036, doi:10.1002/2014JD022493.) Thus, we think it is reasonable to include HIPPO data in the upper troposphere-lower stratosphere region as part of the evaluation of the inversion results. We have added a note in the text justifying the inclusion of the UTLS data.

Line 642-643: It seems that the figure and explanation sentences do not match, so more detailed explanation is needed.
We think the source of confusion for the reviewer is perhaps our lack of a definition for "South America", so we now specify in the text which regions in Fig. 11 correspond to "South America and Africa": "Trop Am", "Temp S Am", "N Africa", and "S Africa".

Line 644: The larger anti-correlations is visible among land area (Bor. N. America and Temp. N. America, Trop. America and Temp. S. America, N. Africa and S. Africa, Temp. Asia and Trop. Asia). The authors should discuss such anti-correlations.
That is a keen observation (the GOSAT inversion exhibits negative correlations of larger magnitude than the in situ inversion for those regions). Note that the point we were trying to make was that "There are a larger *number* of sizable correlations between land regions in the in situ inversion than in the GOSAT inversion." But to provide a more complete discussion, we now add a comment about the larger magnitude of the GOSAT correlations over land, speculating that although GOSAT observations are of higher density over many regions, the column averages tend to reflect mixtures of air from a broader source region than for surface observations, and may thus result in larger error correlations for immediately adjacent regions.

Line 756: A decreased sink in parts of North America (Eurasia) almost matches high temperature anomaly area at 2010 summer. The authors should mention this point in this paragraph. Amazonica aircraft data also could constrain tropical America CO2 flux. https://www.ncdc.noaa.gov/sotc/global/201007
We had discussed the heat waves and droughts of summer 2010 in North America (and Eurasia) in the previous paragraph (citing Guerlet et al., 2013). So it would be repetitive to mention the temperature anomaly in this paragraph.
As for Tropical America, please see our response to the general comments above, including the point about the Amazonica data being available for only a part of our analysis period (namely 2010, so they provide no information on 2010-2009 differences).

Line 767: Measurements from the JR-STATION are significantly important to constrain Eurasia CO2 flux. The authors should include these data to inversion.
Please see our response to the general comments above.

Line 860: The author should consider possibility of using CONTRAIL dataset.

We did suggest that various aircraft data sets could be used as input, although we did not list specific ones besides Amazonica. We now specifically mention CONTRAIL, as well as NOAA's regular aircraft profiles over mostly North America.

Line 863: Satellite retrieval bias also reflected in this paragraph.
This is the same issue as for Line 38; please see our response above.

Line 889-891: GOSAT TIR retrieval also could provide high latitude winter observation. The authors should mention it.
OK, we have added the following sentence after those lines: "Ongoing development of thermal IR (TIR) CO_2 retrievals for GOSAT and the future GOSAT-2 with sensitivity to several layers from the lower troposphere to the lower stratosphere shows promise for producing sufficiently accurate data that could also help to fill NIR retrieval coverage gaps (Saitoh et al., 2017a; b)."

Figure 1: The authors should show validation sites (JR stations, HIPPO and amazonica aircraft).
We have added the JR-STATION and Amazonica sites to the figure (with a different symbol color to represent validation sites). We decided not to show the multiple, irregular HIPPO flight paths for each of two missions on this figure, since they would make the figure rather crowded. We believe that our description of the HIPPO sampling region in the text should be sufficient, and that readers can refer to the HIPPO mission description paper we cited for details and figures.

Figure 5: The authors should show Tropical America to discuss validation against amazon aircraft data.
We appreciate that suggestion. We had carefully selected only three TransCom regions to display in that figure based on space considerations, so we'd prefer not to add another panel. Furthermore, for the purpose of evaluating the model against the aircraft data, it is more illuminating to provide information on both the model and the observations in a plot, as we've done in Fig. 4 and S1, than to show only the model fluxes as in Fig. 5.

Figure 6: The number of observation sites should be shown in the bottom of the figures to know how much the region was constrained.
Regional fluxes are constrained by sites both within and outside of the region. A more appropriate measure of constraint are the uncertainty reductions, which are shown in Table 1.

Figure 10 and 13: The authors should remove or mention outlier (stratospheric observation data?) in the figure. The authors should show from (a) to (f) in the figure.
We assume the reviewer is referring to the outliers in panels (c) and (f), which show the upper altitude bin. We don't see a need to remove outliers, given that we expect our transport model to simulate the tropopause region reasonably well (see response for lines 586-588 above). Also, those particular outliers do not affect our main conclusions about the relative performance of the GOSAT and in situ inversions in that region.
We now add the labels (a) to (f).

Figure 11: The GOSAT inversion seems to enhance dipole phenomena comparing with the in situ inversion. The authors should explain it.
Please see our response for line 644 above.

Anonymous Referee #3

The authors analyze the first year and a quarter of GOSAT column CO2 data (June 2009 - Sept 2010) with a Bayesian synthesis inversion approach, comparing it against a similar inversion of surface in situ CO2 measurements as well as to independent data from the JR-STATION network over Siberia, from the HIPPO transects, and from partial-column CO2 profiles over the Amazon. In the Bayesian synthesis approach, fluxes are estimated across 8-day spans for 108 pre-defined flux regions (obtained by sub-dividing the 22 TransCom3 regions) across late March 2009 through the end of September 2010; the flux patterns assumed inside of each region/span are taken to be the absolute value of the prior fluxes; transport is given by the PCTM off-line model run at 2.0x2.5 deg resolution (lat/lon) with 56 vertical layers. A key advantage of the Bayesian inversion is that a full-rank covariance matrix is obtained for this discretization, providing accurate estimation errors and correlations for analysis purposes. A disadvantage is that flux patterns inside each 8-day span and region cannot be optimized, leading to possible representation errors.

The paper, though long, provides a clear and careful analysis of this initial period of the GOSAT data, attempting to tease apart the influence of errors in the ACOS v3.4 retrievals used from the true flux signals of interest. I believe it is a useful addition to the existing GOSAT literature and should be published here after a few points of clarification (listed below) are addressed.

The main weakness of the work here, in my view, is that the measurement span addressed is quite short and the influence of errors in the initial conditions are likely to be significant further into the flux analysis span than the 40 days at the beginning of the span that have been discarded here. In particular, the June-August 2009 period used to analyze the impact of the 2010 climate drivers for the northern land regions may be feeling the effects of these spin-up errors, since the inversion span begins on March 22, 2009. It is true that the authors attempt to correct for errors in the initial condition by solving for two scalars (a multiple of the initial pattern and an offset) and this may work well, but I would have liked to have seen some sort of sensitivity study addressing the impact of the initial conditions. The comparison to the JR-STATION data suggests that this impact could be substantial. Also, the 2 PgC/year difference in the global flux total estimated by the GOSAT-only inversion in comparison with the in situ inversion clearly points to the impact of the short inversion span: it might have been better to add an additional constraint on the total (land+ocean) flux solved for in the inversion to prevent this difference, since the data themselves do not contain enough trend information to constrain the total.

Overall, though, a lot of good analysis is presented here. I have made some suggestions below for clarifying certain points in the text.

We thank the reviewer for the insightful comments and positive overall assessment. Regarding spin-up, we did put much effort into ensuring a minimal impact of errors in the initial conditions on posterior fluxes after about April 2009, including examining the results of various sensitivity tests. We do mention our conclusion from these tests in Section 2.4, although we had decided not to include much detail on them. But to address the reviewer's concerns, we now add the following details: "For example, for an in situ inversion in which we did not allow adjustments in the i.c. and offset parameters, 8-day average flux results are very similar to those of the baseline inversion, especially after the first two months, with a mean correlation coefficient of 0.95 from June 2009 onward across all TC3 regions and a mean difference of 0.03 Pg C/yr." Discrepancies between the GOSAT posterior concentrations and the JR-STATION

data reflect errors in the GOSAT data and not in the initial conditions, as explained below in the response for lines 767-799.

Regarding a constraint on the total flux, the only information on this available is from atmospheric mole fraction data. We already use most of the long-term surface monitoring sites available in the in situ inversion. And it wouldn't be appropriate to apply a long-term (e.g. several years) average atmospheric growth rate as an inversion constraint, since we are solving for fluxes specific to our analysis period. Also, note that GOSAT observations actually have the advantage of seeing the entire atmospheric column, whereas the in situ observations normally used to estimate the atmospheric growth rate see only the surface or lower troposphere in parts of the world, and thus may not give an accurate estimate for the atmosphere as a whole over a period shorter than the global atmospheric mixing time. On the other hand, GOSAT observations are hampered by the separate problem of retrieval biases. The in situ + GOSAT inversion presumably has the best constrained total flux (though still possibly influenced by GOSAT biases), so our inclusion of those results in the paper gives some indication of the robustness of regional flux results.

Detailed comments:
line 36: Add "Northern Hemisphere" before "high-latitude ocean"? Or do you mean that this applies in the south, as well?
Yes, we do mean both the north and the south, as GOSAT ocean glint observations do not extend into high latitudes in either hemisphere. We have added “northern and southern” before “high-latitude ocean” for clarification.

72: Add "land" before "vegetation"?
OK. (We assume this refers to line *62*.)

96: There is no "Chevallier et al. (2014)" in the reference list. Should the data on this reference be 2013?
We incorrectly wrote “2013” in the reference list. We’ve changed it to “2014”.

112: After "exact solution" add "of the linear equations relating the targeted flux variables to the measurements"? Because the fluxes have been discretized at a fairly coarse spatial and temporal resolution, the approach here does not give an exact solution for the fluxes at fine scales, but it does do so at the coarser resolution targeted here, given the assumed shape of the flux patterns corresponding to each basis function used in the inversion.
Solving for flux scale factors at a coarse resolution actually isn't specific to our batch inversion technique--some of the Kalman filter/smoother methods in use have similar or even coarser spatial resolution (64-156 regions) and temporal resolution (monthly to weekly). Since we are introducing the advantages and disadvantages of the batch technique here, we don’t think there’s a need to mention the flux resolution.

120-122: A downside of this short span is that much of it may be corrupted by spin-up errors, which may last many months after the start of the inversion span.
Please see our response to the overall/general comments above.

184-185: For clarity, replace "...from fossil and biospheric gases" with "from the oxidation of non-CO2 gases from fossil fuel and biospheric burning"?
We have added “fuel” after “fossil” as suggested. For conciseness, we leave out “oxidation” since it is used in the previous sentence and can be inferred from “chemical production” in the

current sentence. We also omit “non-CO2” since it can be inferred from context. “Biospheric burning” is inaccurate as we mention both biogenic and biomass burning gases in the previous sentence. (Examples of biogenic gases include wetland methane and isoprene from vegetation.)

198-200: For the NOAA in situ data, you should give the specific ObsPack file name (which includes the version number) if it came from an ObsPack file, or something equivalent if from some other source.
We did not use an ObsPack file, but rather, the original data files for each individual site. We give the citations recommended by the data providers, namely “Dlugokencky et al., 2013; Andrews et al., 2009”, and in the full references we include the version dates.

226: "and apply a minimum value of 0.01 ppm": it is not clear what this phrase indicates. If there is already an error of 0.3 ppm for the first of up to two possible samples, why is there a need for an additiona 0.01 ppm?
The minimum of 0.01 ppm is needed to avoid standard deviations of 0 for multiple samples (up to “two pairs”, or four samples). We now add the explanation: “to avoid uncertainty values of 0”.

232: Please indicate before this which measurements come at this 30-second frequency – the Japanese continuous sites?
We indeed failed to indicate that. We now add the phrase “30-second-average continuous” early on when we discuss NOAA data flags in the previous paragraph. (We had already specified “hourly” in the context of the Japanese (JMA) data flags.)

251: add a comma before "other"?
OK, done.

252: Add "all" before "GOSAT"?
OK, done.

309-310: Please give the exact equations that implement what you have described here in words. This is needed, because there are different ways to implement what you describe, and these differences can matter to the inversion.
OK, we now provide the equation corresponding to the words, namely Eq. 15 in the paper by Connor et al. (2008), which we cited in that sentence.

374: Add "assumed" before "well-mixed"? For the time after 13 months, was the pattern obtained at the end of the 13 months used in the Jacobian, or was the completely-mixed value used?
This is a statement rather than an assumption, as the detail we provided after that, “(within a range of 0.01 ppm)”, is a quantity based on actual transport model runs. After 13 months, we repeated the pattern at the end of the 13 months until the end of the analysis period.

377-380: It would be useful to describe this SVD procedure in more detail, since discarding singular vectors can completely remove corrections to certain regions at certain times. It might be useful to plot the projection of the singular vectors retained in the fit onto the regions so that the reader can see where the corrections to the prior fluxes are possible and where they are not. What fraction of the original singular value spectrum is truncated and what is retained? Also, usually if one can take the SVD of a matrix,

not much more work is required to obtain the full solution: it is not clear how using an SVD approach helps you deal with the large matrix. Please explain this more. What aspect of your SVD approach allows you to handle the otherwise too-large Jacobian successfully?
The SVD technique avoids the inefficient and numerically unstable inversion of a large Jacobian matrix involved in the direct computation of the Bayesian solution. This is described in more detail in the standard inversion literature, including by Rayner et al. (1999), whom we cite. We actually kept all of the singular vectors in the results presented in the paper; as noted by Rayner et al., the SVD technique avoids amplifying the contributions of small singular values, and thus eliminates the need to set a truncation threshold. We now provide more explanation in the text: "A singular value decomposition (SVD) approach is used *instead of direct computation of Eq. 3 and Eq. 4* to obtain a stable inversion solution *without any need for truncation of singular values below a certain threshold* (Rayner et al., 1999)…."

389-390: "gives mean differences not as close to 0 as in the comparison with the assimilated data": They are actually closer to zero for the in situ inversions, but, yes, quite a bit farther from zero for the GOSAT inversions.
We were actually referring to a comparison of Fig. 3b with 3d (not with 3a) and Fig. 3c with 3a (not with 3d). We acknowledge that the original text was ambiguous; we now add the clarifying phrase "(Fig. 3d and 3a, respectively)" after "not as close to 0 as in the comparison with the assimilated data".

392: "...and have independent random errors": how does the fact that the fit to data not used in the inversions is worse allow you to say that the errors are independent?
Our explanation was indeed flawed. We were comparing the posterior differences with those for the prior model when we stated: "standard deviations that are larger than for the prior" in lines 390-391. So a relevant fact is not that the in situ and GOSAT observations are two different data sets with independent errors, but that the prior model generally exhibits less variability than the posterior model, which has assimilated mole fraction observations, due to random instrument/retrieval errors in the observations, the small sample footprint of the observations compared to the model grid and thus larger variability, and noise created by the inversion process (via error correlations). Thus, posterior model-observation differences would be expected to have a larger standard deviation (when the observations are different from the ones assimilated in the model). We now replace the phrase "have independent random errors" with "combine to produce larger standard deviations than with the less variable prior model, which has not assimilated any data".

430: "fractional": it is not clear here what you mean by this – clarify?
We mean fractional (or percentage) as opposed to absolute (e.g. in Pg C y^{-1}). Smaller regions may have large fractional uncertainties despite having relatively small absolute uncertainties. We have added "(percentage)" after "fractional" for clarity.

432-433: "accounting for error correlations": since you are just aggregating means instead of uncertainties, it is not clear why you need to worry about error correlations – why do you mention it?
We are aggregating "results" here, which include both means and uncertainties.

434-439: Since you have the exact covariances for each region, you could aggregate these (accounting for correlations) and get a posteriori uncertainties for these larger regions. Then you could compare the observed variability to these to see whether

random estimation uncertainties do indeed account for this variability or not. Doing this would be better than just speculating, as you do now.
We did aggregate the exact covariances up to larger regions (perhaps see the response before this one for lines 432-433). Also, the estimation uncertainty for a particular 8-day mean is not really equivalent to an estimate of the variability across different 8-day periods. Finally, we do go beyond speculation, as we discuss the calculated degrees of freedom for signal and noise in that paragraph.

443-447: It is important that you mention here that the observed variability could also be due to systematic errors in either the measurements (especially for the GOSAT case) or in the transport model (especially for the in situ case). By computing the expected random error from your a posteriori covariance matrix, you could potentially rule out random error as the cause, allowing you to attribute the new variability to either a real flux signal or to systematic errors. This is a key reason why you should use your covariance matrix calculation in this analysis.
We did compute posterior uncertainties for aggregated regions, please see the two responses before this one. We also attributed the variability to either signal or noise, as explained in the response before this one. Systematic errors are not as relevant here, since we are discussing the amount of fluctuation from 8-day period to 8-day period rather than biases sustained over longer periods.

446: The dipole behavior mentioned here, in particular, would be reflected in the covariance matrix, if that is in fact the cause of much of the variability.
As described in the three responses before this one, we did utilize the covariance matrices. And as stated in line 447, we discuss dipole behavior associated with negative error correlations later in the paper. Also, note that the numbers of degrees of freedom for signal and noise reported in this paragraph are calculated using the posterior covariance matrix, in a manner described by Rodgers (2000), whom we cite.

474: It might be helpful to mention here in the text that you are comparing your in situ results (not GOSAT results) to CarbonTracker, which also uses only in situ measurements.
We do actually mention that we are comparing our in situ-only inversion with CarbonTracker in this paragraph, in lines 489-490.

489: The sentence starting with "Results" could perhaps be deleted to save space, as it repeats the first sentence of the paragraph.
We appreciate this suggestion. However, we intentionally separated the information about Fig. 7 into the two sentences, since we did not want to include too much detail about the figure in the very first sentence of the paragraph. The two sentences contain distinct information, as the second one specifies that we are comparing only our in situ inversion with CarbonTracker (we had just described the characteristics of the CarbonTracker system in the last few sentences) and that we are showing large regions (aggregated from TransCom regions).

499-502: This is another place where the text could be compacted somewhat – it seems repetitive.
We have now tightened the text here to make it sound less repetitive.

Fig.8 Caption and elsewhere: To avoid having to use the "NEP (x-1)" phraseology everywhere, why not just say you are solving for NEE (which is approximately equal to -NEP)?

NEE includes other fluxes such as emissions from fires and other disturbances; since we prescribed fire emissions and only optimized NEP in the inversions, we needed to be precise in terminology when showing only the NEP component in figures, e.g. Fig. 5 but not Fig. 8.

515-516: "Such a large difference ... is plausible": what evidence can you give to back up your assertion? You have pointed to some plausible causes for the difference, but even so, the difference seems larger than expected. Why did Houweling et al get a difference that was an order of magnitude lower for inversions across a similar span?
For the Houweling et al. results, we gave the total fluxes averaged across 8 models. Within those averages is a substantial amount of inter-model variability. For example, Basu et al. (2013), whose model was one of those included in the Houweling et al. study, reported a 12-month global flux that is ~1.3 Pg C/y less negative for their GOSAT inversion than for their in situ inversion (Sep 2009-Aug 2010, RemoTeC GOSAT retrieval). And Chevallier et al. (2014), who reported results from two of the models included by Houweling et al., found *more negative* global fluxes, by up to 1 Pg C/y, in their GOSAT inversions (ACOS and UoL retrievals) than in their in situ inversions, though for a different 12-month period, 2010. Overall, the results show that the global growth rate can differ substantially over a short period based on either GOSAT or surface observations and among different models. We did show in our paper how the total flux from the inversion can be very sensitive to the time period considered for a 12-month time frame. Also, note that GOSAT column measurements may actually sample the atmosphere better for constraining the short-term growth rate than the surface network, but on the other hand, the result may be affected by retrieval biases. To address the reviewer's concern about the smaller difference in the Houweling et al. results, we now replace "though the difference may not be statistically significant" in lines 523-524 with "with a substantial amount of inter-model variability within those averages".

line 538 and Figure 8: You have used the same term, "Southern Ocean", to refer to the true Southern Ocean (as defined, for example, by TransCom3 as everything south of about 45 deg S) as well as the extra-tropical southern oceans (everything south of 23 deg S). I suggest changing what you call the latter area for clarity.
OK, we now use the phrase "extratropical southern oceans" in the text to refer to the latter. (Likewise with "extratropical northern oceans.") In the figures, for compactness we use the phrase "Southern Oceans," with the "s" at the end to distinguish it from the TransCom "Southern Ocean." (For consistency we add "s" at the end of the other ocean aggregations as well.) Note that this makes Fig. 8 consistent with Fig. 7, in which we had already used plural forms.

568, 571, and Fig 10: The text refers to sub-panels of Figure 10 (a-e), but these labels are absent from the actual figure – please add these labels on the figure.
We now add the labels (a) to (f). (Also for Fig. 13, which is now Fig. S3.)

568-570: "Evaluation of the inversions against latitudinal profiles constructed from HIPPO aircraft measurements, which provide additional sampling over the Pacific, indicates an overestimate by the GOSAT inversion relative to HIPPO in parts of the tropics at lower altitudes": my reading of the figure shows only one, maybe two, points from the GOSAT case that are outside of 1 standard deviation of the observations – this certainly does not seem to be a strong feature of the plots, according to my reading of them. It is not until about 40 deg N that the GOSAT results move positive in panel a).
We agree that there is not a general overestimate by the GOSAT inversion in the tropics relative to HIPPO, as there are similar numbers of points higher and lower than HIPPO. We now modify the sentence to read "does not indicate any widespread overestimate by the GOSAT inversion

relative to the observations in the tropics (Fig. 10a-f), unlike what was seen in the comparison with the more globally distributed surface observations."

632: Replace "elaborate on the subject of" with "discuss"? Less wordy...
We think "elaborate" is a more precise term here, but we do shorten the phrase to "elaborate on".

695-696: "accounting for the riverine flux, the 1_ range for the in situ inversion overlaps with that of GCP": I believe you are incorrectly applying the riverine flux correction here. The GCP number of -2.5 should be decreased to -2.0 PgC/yr when turning it from an anthropogenic uptake into a total net (anthropogenic+natural) uptake, since the natural cycle (driven by the riverine fluxes into the ocean at the river mouths) has a net 0.5 PgC/yr outgassing – that outgassing counteracts a corresponding amount of anthropogenic uptake, reducing the total uptake to -2.0 PgC/yr. I.e., -2.5 + (+0.5) = -2.0. Given that, both your GOSAT-only and in situ-only ocean uptakes are still outside the 1 sigma ranges for the GCP number.
We actually did apply the riverine flux correction correctly. Two paragraphs before this one, we explained that "The difference between our inversion estimates and the GCP estimate is actually even larger than suggested by those numbers, given that a background river to ocean flux of ~0.5 Pg C y^{-1} should be subtracted from our ocean flux to make it comparable to the GCP ocean sink, which refers to net uptake of *anthropogenic* CO_2". Rather than converting the GCP number to a net uptake, we just left it as an anthropogenic uptake, and implied that 0.5 Pg C y^{-1} should be subtracted from our ocean fluxes for comparison with GCP. (We did not feel it was necessary to explicitly give the adjusted numbers.) We think the source of misunderstanding is that we were comparing the 1-sigma ranges for our estimates with the 1-sigma range of GCP, rather than our central estimates with the 1-sigma range. To clarify, we have modified the text to: "accounting for the riverine flux, the 1σ range for the in situ inversion overlaps with the 1σ range of GCP, while the 1σ range for the GOSAT inversion is still just outside of that of GCP."

Figure 14: I would suggest using some color other than cyan to depict the tight-prior GOSAT results here. As things stand now, it is much too easy to confuse that case with the in situ-only results on Figure 9. One has to read the caption and legend carefully to see that you have changed what is shown in cyan at the moment.
OK, we have changed the color to green. (The figure is now Fig. S4.)

724: "substantially larger global total budget": it would be clearer to say that the total flux is more positive in the GOSAT case, since "larger" depends on whether the fossil fuel has been added onto the total or not.
OK, we have changed the wording to "substantially more positive global total budget".

734-735: "and an increased source in the tropics of _2 Pg C y-1 in the GOSAT inversion relative to the in situ inversion.": I think that it is important to note that this change from the in situ-only results in the tropics is accompanied by a change in the global total of the same magnitude and sign; in other words, the change is directly related to the fact that the global total is not well-constrained in this short-span inversion. This might be expected to change in an inversion over a longer span, for which the global total is better constrained.
The difference between the in situ and GOSAT budgets in the tropics could indeed be related to insufficient constraints on a 12-month time scale for one or both of them (as we hypothesized for the global budgets), so that the difference could be much smaller when averaged over a

span of several years or longer. However, the in situ + GOSAT inversion gives an indication of the impact of an under-constrained global total flux on regional flux estimates, given that the combined data sets provide more constraint than either of the data sets alone. The in situ + GOSAT inversion produces a global flux close to mid-way between the in situ-only and GOSAT-only inversions, while it produces a Tropic Land + Ocean flux much closer to that of the GOSAT inversion than to the in situ inversion. This suggests some degree of independence of the GOSAT-inferred regional result from the global result. We now mention this in the manuscript.

767-799: This whole discussion of the Eurasian source in 2010 and the examination of JR-STATION sites suggests to me that the growing season results in 2009 could well be affected by spin-up issues in the inversion. That could explain why the GOSAT inversion results agree with the data at VGN, AZV, and KRS in 2010, but are too negative in 2009. If that is the explanation, the agreement with the Guerlet (2013) result would be more due to that modeling issue, rather than any real climate-related driver.
Please see our response to the general comments above regarding spin-up issues. In addition, it is important to be aware that inaccurate initial conditions do not affect the fit of the posterior concentrations to observations, but rather, the correctness of the posterior fluxes. The inversion optimizes the fit to observations through flux adjustments regardless of the i.c., but there would be errors in the final fluxes if the i.c. were incorrect (after adjustment of the i.c. and offset parameters in the case of our inversion setup). Thus, any discrepancies between our GOSAT posterior concentrations and the JR-STATION data would reflect errors in the GOSAT data and not errors in the i.c.

899-901: "Thus, it may not be accurate to assume that year-to-year posterior flux differences are insensitive to satellite retrieval biases, as was done in the other study."
This would be a good place to note that spin-up errors in this study (as well as the Houweling study) could also be adversely affecting the 2009 flux results, as well as the 2010-2009 shift.
Please see our response to the previous comment (on lines 767-799).

A Global Synthesis Inversion Analysis of Recent Variability in CO_2 Fluxes Using GOSAT and In Situ Observations

James S. Wang,[1,2] S. Randolph Kawa,[2] G. James Collatz,[2] Motoki Sasakawa,[3] Luciana V. Gatti,[4] Toshinobu Machida,[3] Yuping Liu,[5,2] and Michael E. Manyin[5,2]

[1]Universities Space Research Association, Columbia, MD, USA, james.s.wang@nasa.gov

[2]NASA Goddard Space Flight Center, Greenbelt, MD, USA

[3]National Institute for Environmental Studies, Center for Global Environmental Research, Ibaraki, Tsukuba Onogawa, Japan

[4]Instituto de Pesquisas Energéticas e Nucleares (IPEN)–Comissao Nacional de Energia Nuclear (CNEN), Sao Paulo, Brazil

[5]Science Systems and Applications, Inc., Lanham, MD, USA

Submitted to Atmospheric Chemistry and Physics (as part of Special Issue "The 10th International Carbon Dioxide Conference (ICDC10) and the 19th WMO/IAEA Meeting on Carbon Dioxide, other Greenhouse Gases and Related Measurement Techniques (GGMT-2017)")

Abstract

The precise contribution of the two major sinks for anthropogenic CO_2 emissions, terrestrial vegetation and the ocean, and their location and year-to-year variability are not well understood. Top-down estimates of the spatiotemporal variations in emissions and uptake of CO_2 are expected to benefit from the increasing measurement density brought by recent in situ and remote CO_2 observations. We uniquely apply a batch Bayesian synthesis inversion at relatively high resolution to in situ surface observations and bias-corrected GOSAT satellite column CO_2 retrievals to deduce the global distributions of natural CO_2 fluxes during 2009-2010. Our objectives include evaluating bottom-up prior flux estimates, assessing the value added by the satellite data, and examining the impacts of inversion technique and assumptions on posterior fluxes and uncertainties. The GOSAT inversion is generally better constrained than the in situ inversion, with smaller posterior regional flux uncertainties and correlations, because of greater spatial coverage, except over North America and northern and southern high-latitude ocean. Complementarity of the in situ and GOSAT data enhances uncertainty reductions in a joint inversion; however, remaining coverage gaps, including those associated with spatial and temporal ~~gaps in~~ sampling biases in the passive satellite measurements, still limit the ability to accurately resolve fluxes down to the sub-continental/sub-ocean basin scale. The GOSAT inversion produces a shift in the global CO_2 sink from the tropics to the north and south relative to the prior, and an increased source in the tropics of ~2 Pg C y^{-1} relative to the in situ inversion, similar to what is seen in studies using other inversion approaches. This result may be driven by sampling and residual retrieval biases in the GOSAT data, as suggested by significant discrepancies between posterior CO_2 distributions and surface in situ and HIPPO mission aircraft data. While the shift in the global sink appears to be a robust feature of the inversions, the

47 partitioning of the sink between land and ocean in the inversions using either in situ or GOSAT
48 data is found to be sensitive to prior uncertainties because of negative correlations in the flux
49 errors. The GOSAT inversion indicates significantly less CO_2 uptake in summer of 2010 than in
50 2009 across northern regions, consistent with the impact of observed severe heat waves and
51 drought. However, observations from an in situ network in Siberia imply that the GOSAT
52 inversion exaggerates the 2010-2009 difference in uptake in that region, while the prior CASA-
53 GFED model of net ecosystem production and fire emissions reasonably estimates that quantity.
54 The prior, in situ posterior, and GOSAT posterior all indicate greater uptake over North America
55 in spring to early summer of 2010 than in 2009, consistent with wetter conditions. The GOSAT
56 inversion does not show the expected impact on fluxes of a 2010 drought in the Amazon;
57 evaluation of posterior mole fractions against local aircraft profiles suggests that time-varying
58 GOSAT coverage can bias estimation of flux interannual variability in this region.
59

1. Introduction

About one-half of the global CO_2 emissions from fossil fuel combustion and deforestation accumulates in the atmosphere (Le Quéré et al., 2015), where it contributes to global climate change. The rest is taken up by land vegetation and the ocean. The precise contribution of the two sinks, their location and year-to-year variability, and the environmental controls on the variability are, however, not well understood. Top-down methods involving atmospheric inverse modeling have been used extensively to quantify natural CO_2 fluxes (e.g. Enting and Mansbridge, 1989; Ciais et al., 2010). An advantage of this approach over bottom-up methods such as forest inventories (Pan et al., 2011; Hayes et al., 2012) or direct flux measurements (Baldocchi et al., 2001; Chevallier et al., 2012) is that measurements of atmospheric CO_2 mole fractions generally contain the influence of fluxes over a spatial scale substantially larger than that of individual forest plots or flux measurements, so that errors from extrapolating measurements to climatically relevant scales (e.g. ecosystem, sub-continental, or global) are mitigated. However, the accuracy of top-down methods is limited by ~~sparse~~ incomplete data coverage (especially for highly precise but sparse in situ observation networks), uncertainties in atmospheric transport modeling, and mixing of signals from different flux types such as anthropogenic and natural.

With the advent of retrievals of atmospheric CO_2 mole fraction from satellites, including the Japanese Greenhouse gases Observing SATellite (GOSAT) (Yokota et al., 2009) and the NASA Orbiting Carbon Observatory-2 (OCO-2) (Crisp, 2015; Eldering et al., 2017), data coverage has improved ~~greatly~~substantially. Making measurements since 2009, GOSAT is the first satellite in orbit designed specifically to measure column mixing ratios of CO_2 (as well as methane) with substantial sensitivity to the lower troposphere, close to surface fluxes. A number

of modeling groups have conducted CO_2 flux inversions using synthetic GOSAT data (Liu et al., 2014) and actual data (Takagi et al., 2011; Maksyutov et al., 2013; Basu et al., 2013; Saeki et al., 2013a; Deng et al., 2014; Chevallier et al., 2014; Takagi et al., 2014; Reuter et al., 2014; Houweling et al., 2015; Deng et al., 2016). Unlike in situ measurements, which are calibrated directly for the gas of interest, remote sensing involves challenges in precision and accuracy stemming from the measuring of radiance. The retrievals rely on modeling of radiative transfer involving complicated absorption and scattering by the atmosphere and reflection from the surface (e.g. Connor et al., 2008; O'Dell et al., 2012). Passive measurements that rely on reflected sunlight are more prone to errors than active measurements, as they are affected by not only errors related to meteorological parameters and instrument noise but also systematic errors related to scattering by clouds and aerosols, which can dominate the error budget (Kawa et al., 2010; O'Dell et al., 2012). Furthermore, passive measurements have coverage gaps where there is insufficient sunlight and where there is excessive scattering.

In addition to the model transport examined by a number of inversion intercomparison studies (e.g. Gurney et al., 2002; Baker et al. 2006), the inversion technique and assumptions can contribute to substantial differences in results. For example, Chevallier et al. (2014) found that significant differences in hemispheric and regional flux estimates can stem from differences in Bayesian inversion techniques, transport models, a priori flux estimates, and satellite CO_2 retrievals. Houweling et al. (2015) presented an intercomparison of 8 different inversions using 5 independent GOSAT retrievals, and also found substantial differences in optimized fluxes at the regional level, with modeling differences (priors, transport, inversion technique) contributing approximately as much to the spread in results on land as the different satellite retrievals used.

105 In this paper, we present inversions of GOSAT and in situ data using a distinct technique,
106 which are compared with results from other studies. All of the previous GOSAT satellite data
107 inversions have used computationally-efficient approaches, such as variational and ensemble
108 Kalman filter data assimilation, to handle the large amounts of data generated by satellites and
109 the relatively large number of flux regions whose estimation is enabled by such data. The
110 computational efficiency of these approaches results from numerical approximations. In this
111 study, we apply a traditional, batch, Bayesian synthesis inversion approach (e.g. Baker et al.,
112 2006) at ~~relatively~~ high spatiotemporal resolution relative to previous batch inversions to
113 estimate global, interannually varying CO_2 fluxes from satellite and in situ data. Advantages of
114 this technique include generation of an exact solution along with a full-rank error covariance
115 matrix (e.g. Chatterjee and Michalak, 2013), and an unlimited time window during which fluxes
116 may influence observations, unlike the limits typically imposed in Kalman filter techniques. The
117 major disadvantages of the batch technique are that computational requirements limit the
118 spatiotemporal resolution at which the inversion can be solved and the size of the data set that
119 can be ingested, a large number of transport model runs is required to pre-compute the basis
120 functions (i.e. Jacobian matrix), and the handling of the resulting volume of model output is very
121 time-consuming at relatively high resolution.
122 We estimate natural terrestrial and oceanic fluxes over the period May 2009 through
123 September 2010. The analysis spans two full boreal summers; longer periods were prohibited by
124 the computational effort. The objectives of this study are: 1) to understand recent variability of
125 the global carbon cycle, 2) to evaluate the bottom-up flux estimates used for the priors, 3) to
126 compare fluxes and uncertainties inferred using in situ observations, GOSAT observations, and
127 the two data sets combined and to assess the value added by the satellite data, and 4) to generate

inversion results using a unique Bayesian inversion technique for comparison with other approaches.

Section 2 provides details on the inputs and inversion methods. Section 3 presents prior and posterior model CO_2 mole fractions and their evaluation against independent data sets, fluxes and uncertainties at various spatial and temporal scales, and comparisons with results from inversions conducted by other groups. We discuss the robustness of results, and examine in particular their sensitivity to assumed prior flux uncertainties. We then analyze the possible impacts of several climatic events during the analysis period on CO_2 fluxes. Section 4 contains concluding remarks.

2. Methods

Our method is based on that used in the TransCom 3 (TC3) CO_2 inversion intercomparisons (Gurney et al., 2002; Baker et al. 2006) and that of Butler et al. (2010), the latter representing an advance over the TC3 method in that they accounted for interannual variations in transport and optimized fluxes at a higher spatial resolution. Our method involves further advances over that of Butler et al. (2010), including higher spatial and temporal resolution for the optimized fluxes, and the use of individual flask-air observations and daily averages for continuous observations rather than monthly averages. Inversion theoretical studies and intercomparisons have suggested that coarse resolution for flux optimization can produce biased estimates, i.e. estimates that suffer from aggregation error (Kaminski et al., 2001; Engelen et al., 2002; Gourdji et al., 2012). Although observation networks may not necessarily provide sufficient constraints on fluxes at high resolutions, Gourdji et al. (2012) adopted the approach of

151 estimating fluxes first at fine scales and then aggregating to better-constrained resolutions to
152 minimize aggregation errors. The high spatiotemporal resolution of our inversion relative to
153 other global batch ~~Bayesian synthesis~~ inversions would be expected to reduce aggregation errors.
154 Similarly, use of higher temporal resolution observations allows our inversion to more precisely
155 capture variability due to transport and thus more accurately estimate fluxes. Details on our
156 inversion methodology are provided in the sub-sections below.
157
158 2.1. A priori fluxes and uncertainties
159 Prior estimates for net ecosystem production (NEP = photosynthesis − respiration) and
160 fire emissions (wildfires, biomass burning, and biofuel burning) come from the Carnegie-Ames-
161 Stanford-Approach (CASA) biogeochemical model coupled to version 3 of the Global Fire
162 Emissions Database (GFED3) (Randerson et al., 1996; van der Werf et al., 2006; 2010). CASA-
163 GFED is driven with data on fraction of absorbed photosynthetically active radiation (FPAR)
164 derived from the AVHRR satellite series (Pinzon et al, 2014; Los et al., 2000), burned area from
165 MODIS (Giglio et al, 2010), and meteorology (precipitation, temperature, and solar radiation)
166 from the Modern-Era Retrospective Analysis for Research and Applications (MERRA)
167 (Rienecker et al., 2011). CASA-GFED fluxes are generated at 0.5° x 0.5° resolution. For use in
168 the atmospheric transport model, monthly fluxes are downscaled to 3-hourly values using solar
169 radiation and temperature (Olsen and Randerson, 2004) along with MODIS 8-day satellite fire
170 detections (Giglio et al., 2006). In general, the biosphere is close to neutral in the CASA-GFED
171 simulation, i.e. there is no long-term net sink although there can be interannual variations in the
172 balance between uptake and release. In the version of CASA used here, a sink of ~100 Tg C y^{-1}
173 is induced by crop harvest in the U.S. Midwest that is prescribed based on National Agriculture

Statistics Service data on crop area and harvest. Although respiration of the harvested products is neglected, the underestimate of emissions that is implied is geographically dispersed and in principle correctable by the inversion.

For air-sea CO_2 exchange, monthly, climatological, measurement-based fluxes are taken from Takahashi et al. (2009) for the reference year 2000 on a 4° x 5° lat/lon grid. In contrast to the CASA-GFED flux being close to neutral on a global basis, the prior ocean flux forms a net sink of 1.4 Pg C y^{-1}. (Note that the uncertainties we assign to the prior fluxes, discussed below, are large enough to accommodate possible biases, e.g. the neutral biosphere rather than a sizable net land sink as suggested by the literature.) For fossil CO_2, 1° x 1°, monthly- and interannually-varying emissions are taken from the Carbon Dioxide Information Analysis Center (CDIAC) inventory (Andres et al., 2012). This includes CO_2 from cement production but not international shipping and aviation emissions. Oxidation of reduced carbon-containing gases from fossil fuels in the atmosphere (~5% of the emissions; Nassar et al., 2010) is neglected, and the entire amount of the emissions is released as CO_2 at the surface. Similarly, CO_2 from oxidation of biogenic and biomass burning gases is neglected. The total amount of CO_2 chemical production from fossil-fuel and biospheric gases is estimated to be ~1 Pg C y^{-1} (for year 2006; Nassar et al., 2010).

A priori flux uncertainties are derived from those assumed in the TC3 studies (Table 1), rescaled to our smaller regions and shorter periods with the same approach as Feng et al. (2009). A priori spatial and temporal error correlations are neglected in our standard inversions. The neglect of a priori spatial error correlations is justified by the size of our flux optimization regions, with dimensions on the order of one thousand to several thousand km, likely greater than the error correlation lengths for our 2° × 2.5° grid-level fluxes. For example, Chevallier et

196 al. (2012) estimated a correlation e-folding length of ~500 km for a grid size close to ours of 300
197 km × 300 km based on comparison of a terrestrial ecosystem model with global flux tower data.
198
199 2.2. Observations and uncertainties
200 For constraining fluxes at relatively high temporal resolution, observations are chosen
201 that consist of discrete whole-air samples collected in glass flasks approximately weekly and
202 continuous in situ tall tower measurements of CO_2 mole fraction from the NOAA ESRL Carbon
203 Cycle Cooperative Global Air Sampling Network (Dlugokencky et al., 2013; Andrews et al.,
204 2009) supplemented with continuous ground-based measurements at 3 sites in East Asia from the
205 Japan Meteorological Agency (JMA) network (http://ds.data.jma.go.jp/gmd/wdcgg/cgi-
206 bin/wdcgg/catalogue.cgi, accessed 14 Mar 2013; Tsutsumi et al., 2006). Both data sets are
207 calibrated to the WMO-X2007 scale. In the present study, these data sets are referred to
208 collectively as "in situ" observations. The 87 sites (Fig. 1a) are chosen based on data availability
209 for the analysis period, Mar 2009-Sep 2010. Individual flask-air observations are used in the
210 inversions (with the average taken where there are multiple measurements at a particular hour—
211 up to two pairs of duplicate samples), and for the continuous measurements, afternoon averages
212 are used (between 1200 and 1700 local standard time), avoiding the difficulty of simulating the
213 effects of shallow nighttime boundary layers. For the towers, data from the highest level only is
214 used. We apply minimal filtering of the data. For the NOAA data sets, we exclude only the
215 flask samples or 30-second-average continuous data with "rejection" flags, retaining data with
216 "selection" flags (NOAA uses statistical filters and other information such as wind direction to
217 flag data that are likely valid but do not meet certain criteria such as being representative of well-
218 mixed, background conditions), since the ~~relatively~~ reasonably high-resolution transport model

used here (Sect. 2.3) captures much of the variability in the observations beyond background levels. Furthermore, observations strongly influenced by local fluxes are typically assigned larger uncertainties by our scheme (described below), and therefore have less weight in the inversion. For the JMA data, we omit only the hourly data with flag = 0, meaning the number of samples is below a certain level, the standard deviation is high, and there is a large discrepancy with one or both adjacent hourly values. Although some of the observation sites used in our inversion are located close to each other, there is never any exact overlap in grid box (altitude and/or longitude-latitude) or in time. Thus, all of those sites are kept for the inversions, with observations at each site and day treated as independent (i.e. neglecting error correlations).

We estimate the uncertainties for the flask-air observations as the root sum square (RSS) of two uncertainty components: the standard deviation of the observations ~~within~~ at a particular hour ~~(up to two pairs of duplicate samples)~~ and a simple estimate of the model transport/representation error. For the first uncertainty component, we assign a value of 0.3 ppm if there is only one sample, and apply a minimum value of 0.01 ppm to avoid uncertainty values of 0. The transport/representation error estimation is similar to that of the NOAA CarbonTracker (CT) CO_2 data assimilation system (prior to the CT 2015 version) (Peters et al., 2007; http://carbontracker.noaa.gov), whereby a fixed "model-data mismatch" is assigned based on the type of site, e.g. marine, coastal, continental, or polluted, ranging from 0.4 ppm to 4 ppm. For the continuous measurements, we take the RSS of two uncertainty components: the afternoon root mean square (RMS) of the uncertainties of the 30-second or hourly average observations reported by the data providers (divided by the square root of the number of observations, N), and the standard error of all the 30-second/hourly mole fractions within an afternoon period. This represents an attempt to account for instrument error as well as transport/representation error. In

addition, based on initial inversion results, we enlarged all in situ total observation uncertainties by a factor of 2 to lower the normalized ~~chi-squared (χ^2)~~ posterior cost function value (defined in Section 2.4) closer to 1 as appropriate for the chi-squared (χ^2) distribution (the final value of which is shown in Table 2). (Another test showed that further enlargement of the uncertainties to 3 times the original values, while lowering the cost function value further, does not substantially change the posterior fluxes overall.)

GOSAT measures reflected sunlight in a sun-synchronous orbit with a 3-day repeat cycle and a 10.5 km diameter footprint when in nadir mode (Yokota et al., 2009). The spacing between soundings is ~250 km along-track and ~160 km or ~260 km cross-track (for 5-point/3-point sampling before/after Aug 2010). We use the ACOS B3.4 near infrared (NIR) retrieval of column-average CO_2 dry air mole fraction (XCO_2), with data from June 2009 onward (O'Dell et al., 2012; Osterman et al., 2013). Filtered and bias-corrected land nadir, including high (H) gain and medium (M) gain, and ocean glint data are provided. Three truth metrics were used together to correct biases (separately for H gain, M gain, and ocean glint) (Osterman et al., 2013; Lindqvist et al., 2015; Kulawik et al., 2016): 1) an ensemble of transport model simulations optimized against in situ observations, 2) coincident ground-based column observations from the Total Carbon Column Observing Network (TCCON), which are calibrated to aircraft in situ profiles linked to the WMO scale (Wunch et al., 2011), and 3) the assumption that CO_2 mole fraction ought to exhibit little spatiotemporal variability in the Southern Hemisphere mid-latitudes, other than a seasonal cycle and long-term trend. For our inversions, we use the average of all GOSAT observations falling within a given 2° latitude × 2.5° longitude transport model column in a given hour. Figure 1b shows the frequency of the ACOS GOSAT observations across the model grid.

The values assumed for the GOSAT uncertainties are based in part on the retrieval uncertainties provided with the ACOS data set. Following guidance from the data providers, these are inflated by a factor of 2 over land and 1.25 over ocean for more realistic estimates of the uncertainties (C. O'Dell, pers. comm., 2013); Kulawik et al. (2016) recommended an overall scale factor of 1.9 for the similar ACOS B3.5 data set. In the case of multiple observations within a model grid, we estimate the overall uncertainty as the RMS of the uncertainties of the individual observations, divided by the square root of N. Error correlations between observations in different model grids and at different hours are neglected.

Inversions are conducted using different combinations of data, including the in situ data ("in situ-only"), the GOSAT data ("GOSAT-only"), and both ("in situ + GOSAT").

We use several additional data sets for independent evaluation of the inversion results. Aircraft measurements from the HIAPER Pole-to-Pole Observations (HIPPO) campaign consist of vertical profiles of climate-relevant gases and aerosols from the surface to as high as the lower stratosphere, spanning a wide range of latitudes mostly over the Pacific region (Wofsy et al., 2011). Five missions were conducted during different seasons in 2009-2011, with two of the missions overlapping with our analysis period. We use the "best available" CO_2 values derived from multiple measurement systems from the merged 10-second data product (Wofsy et al., 2012). Another data set, the 'Amazonica' aircraft measurements over the Amazon basin, is useful for evaluating inversion performance over tropical land. These measurements consist of profiles of several gases including CO_2 determined from flask samples from just above the forest canopy to 4.4 km altitude over 4 sites across the Brazilian Amazon starting in 2010, taken approximately biweekly (Gatti et al., 2014, 2016). Finally, the Japan-Russia Siberian Tall Tower Inland Observation Network (JR-STATION) of towers provides continuous in situ

measurements of CO_2 and CH_4 over different ecosystem types across Siberia beginning in 2002 (Sasakawa et al., 2010; Sasakawa et al., 2013). The JR-STATION data have been used in combination with other in situ observations in CO_2 flux inversions (Saeki et al., 2013b; Kim et al., 2017).

2.3. Atmospheric transport model and model sampling

We use the Parameterized Chemistry and Transport Model (PCTM) (Kawa et al., 2004), with meteorology from the NASA Global Modeling and Assimilation Office (GMAO) MERRA reanalysis (Rienecker et al., 2011). For this analysis, PCTM was run at a resolution of 2° latitude × 2.5° longitude and 56 hybrid terrain-following levels up to 0.4 hPa, and hourly temporal resolution. A "pressure fixer" scheme has been implemented to ensure tracer mass conservation, the lack of which can be a significant problem with assimilated winds (Kawa et al., 2004). Evaluation of PCTM over the years has shown it to a reliable tool for carbon cycle studies. For example, Kawa et al. (2004) showed that the SF_6 distribution from PCTM was consistent with that of observations and of the models in TransCom 2, suggesting that the interhemispheric and vertical transport were reasonable. ~~This version of~~ PCTM performed well in boundary layer turbulent mixing compared to most of ~~was among~~ the other ~~transport~~ models ~~that participated~~ in a TransCom investigation of the CO_2 diurnal cycle~~intercomparisons~~ (~~Gurney et al., 2005;~~ Law et al., 2008). The TransCom-CH_4 intercomparison (Patra et al., 2011) showed that a more recent version of PCTM performed very well relative to observations in its interhemispheric gradients of SF_6, CH_3CCl_3, and CH_4 and interhemispheric exchange time, and follow-on studies (Saito et al., 2013; Belikov et al., 2013) demonstrated through evaluation against observed CH_4 and ^{222}Rn that the convective vertical mixing in PCTM was satisfactory overall.

Offshore prior terrestrial biospheric and fossil fluxes are redistributed to the nearest onshore grid cells in the model grid to counteract diffusion caused by our regridding the original fluxes to the coarser 2° x 2.5° resolution, as recommended in the TC3 protocol (Gurney et al., 2000).

The model is initialized with a concentration field appropriate for March 22, 2009 from a multi-year PCTM run with prior fluxes. The initial conditions are optimized in the inversions, as described in Sect. 2.4.

PCTM is sampled at grid cells containing in situ observation sites or GOSAT soundings, at the hours corresponding to the observations. To mimic the sampling protocol for coastal flask sites, which favors clean, onshore wind conditions, the model is sampled at the neighboring offshore grid cell if the cell containing the site is considered land according to a land/ocean mask. For in situ sites in general, ~~the~~ an appropriate vertical level as well as horizontal location is selected. Specifically, the model CO_2 profile is interpolated to ~~the~~ a level corresponding on average to the altitude above sea level of the observation site. This procedure is relevant primarily for mountain sites and tall towers as well as aircraft samples; the lowest model ~~level~~ layer (with a thickness of ~100 m on average) was used for most other sites.

Model columns are weighted using ACOS column averaging kernels ~~applied to deviations of model CO_2 profiles from ACOS a priori profiles~~, as in the following (Eq. 15 from Connor et al., 2008):

$$\boldsymbol{X^{m}_{CO2}} = \boldsymbol{X^{a}_{CO2}} + \sum_j \mathbf{h}_j \mathbf{a}_{\mathbf{CO2},j} (\mathbf{x}_m - \mathbf{x}_a)_j, \quad (1)$$

where X^m_{CO2} (X^a_{CO2}) refers to the model (ACOS a priori) column average mole fraction, **h** is the pressure weighting function, $\mathbf{a}_{CO2}$ is the column averaging kernel, **x** refers to a CO_2 profile, and j is the level index.

Time series of model and observed mole fractions at selected flask and continuous sites spanning a range of latitudes, longitudes, elevations, and proximity to major fluxes are shown for the prior and for the in situ-only inversion in Fig. 2. The prior model as well as the in situ inversion captures much of the observed synoptic-scale variability. This suggests that the PCTM transport is reasonably accurate, consistent with the findings of Parazoo et al. (2008) and Law et al. (2008).

2.4. Inversion approach

The batch, Bayesian synthesis inversion approach optimizes in a single step the agreement between model and observed CO_2 mole fractions and between a priori and a posteriori flux estimates in a least-squares manner (e.g. Enting et al., 1995). As in the paper by Baker et al. (2006), the cost function minimized in this approach can be expressed as

$$J = (\mathbf{c}_{\text{obs}} - \mathbf{c}_{\text{fwd}} - \mathbf{Hx})^T \mathbf{R}^{-1} (\mathbf{c}_{\text{obs}} - \mathbf{c}_{\text{fwd}} - \mathbf{Hx}) + (\mathbf{x}_0 - \mathbf{x})^T \mathbf{P}_0^{-1} (\mathbf{x}_0 - \mathbf{x}), \quad (1\text{~~}2)$$

where $\mathbf{c}_{\text{obs}} - \mathbf{c}_{\text{fwd}}$ are mismatches between the observations and the mole fractions produced by the prior fluxes, **H** is the Jacobian matrix relating model mole fractions at the observation locations to regional flux adjustments **x** (note that **x** is used differently here than in Eq. 1), **R** is the covariance matrix for the errors in $\mathbf{c}_{\text{obs}} - \mathbf{c}_{\text{fwd}}$, $\mathbf{x}_0$ is an a priori estimate of the flux

adjustments, and $\mathbf{P}_0$ is the covariance matrix for the errors in $\mathbf{x}_0$. The solution for the a posteriori flux adjustments, $\hat{\mathbf{x}}$, is

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{P}_0^{-1})^{-1}(\mathbf{H}^T\mathbf{R}^{-1}(\mathbf{c}_{\text{obs}} - \mathbf{c}_{\text{fwd}}) + \mathbf{P}_0^{-1}\mathbf{x}_0), \quad (23)$$

and the a posteriori error covariance matrix is given by

$$\mathbf{P} = (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{P}_0^{-1})^{-1}. \quad (34)$$

Importantly, the posterior uncertainties do not account for possible biases, given that the Bayesian inversion framework adopted here, as in other CO_2 studies, assumes Gaussian error distributions with no bias (observation, transport, prior, etc.).

This study focuses on the variability of natural fluxes (terrestrial NEP and ocean), and thus considers adjustments to those fluxes only, assuming the prior estimates for the fossil and fire fluxes are correct. This is commonly done in CO_2 inversion studies (e.g. Gurney et al., 2002; Peters et al., 2007; Basu et al., 2013), with the rationale that the anthropogenic emissions are relatively well known, at least at the coarse spatial scales of most global inversions. In our inversion, flux adjustments are solved for at a resolution of 8 days and for each of 108 regions that are modified from the 144 regions of the Feng et al. (2009) inversion (Fig. 1a), which are in turn subdivided from the TC3 regions. (The choice of an 8-day flux interval is based on data considerations, e.g. the quasi-weekly frequency of the flask measurements and reasonable sampling by GOSAT.) This is a significantly higher resolution than the monthly intervals and 22/47 regions in the previous batch inversions of TC3/Butler et al. (2010), which allows us to

take advantage of the relatively high density of the GOSAT observations. One of our regions consists of low-flux ~~land~~ areas (e.g. Greenland, Antarctica) as well as small offshore areas that contain non-zero terrestrial biospheric fluxes but do not fit into any of the TC3-based land regions, similar to what was done by Feng et al. (2009). We also created a region that includes areas with non-zero oceanic fluxes that do not fit into any of the TC3-based ocean regions according to our gridding scheme.

Grid-scale spatial patterns are imposed in our flux adjustments based on the natural fluxes, similar to TC3 and Butler et al. (2010), except that we use patterns specific to our prior NEP or air-sea flux averaged over each particular 8-day period, rather than annual mean net primary productivity (NPP) patterns over land and spatially constant patterns over the ocean. To ensure net changes in flux are possible across each region, absolute values are used for the flux patterns. Prior values of 0 are specified for all flux adjustments.

The initial conditions (i.c.) are also optimized at the same time as the fluxes via two parameters: a scale factor to the i.c. tracer (described below) that allows for overall adjustment of spatial gradients, and a globally uniform offset. A priori uncertainties of 0.01 for the scale factor and 30 ppm for the offset are prescribed. Inversion results from March 22 through April 30, 2009 are discarded to avoid the influence of any inaccuracies in the ~~initial conditions~~i.c. (Our tests showed that inferred fluxes after the first two months are insensitive to the treatment of i.c. For example, for an in situ inversion in which we did not allow adjustments in the i.c. and offset parameters, 8-day average flux results are very similar to those of the baseline inversion, especially after the first two months, with a mean correlation coefficient of 0.95 from June 2009 onward across all TC3 regions and a mean difference of 0.03 Pg C/yr.) Although the GOSAT data set begins in June 2009, the observations can provide some constraint on earlier fluxes.

For generating the prior mole fractions, $\mathbf{c}_{fwd}$, and constructing the Jacobian matrix, $\mathbf{H}$, transport model runs were performed for each of the prior flux types and an i.c. tracer, as well as a run with a flux pulse (normalized to 1 Pg C y^{-1}) for each of the 108 regions and 71 8-day periods. (The last period in 2009 is shortened to 5 days to fit cleanly within the year.) The i.c. tracer is initialized as described in Sect. 2.3 and transported without emissions or removals for the duration of the analysis period. Each flux pulse is transported for up to 13 months, after which the atmosphere is well mixed (within a range of 0.01 ppm). This procedure generated a massive amount of 3-D model output, ~30 terabytes (compressed). All of the model output was then sampled at the observation locations and times.

A singular value decomposition (SVD) approach is used instead of direct computation of Eq. 3 and Eq. 4 to obtain a stable inversion solution without any need for truncation of singular values below a certain threshold (Rayner et al., 1999). Use of the SVD technique is especially helpful in the case of the inversions using GOSAT data, since the Jacobian matrix is too large (92762 (102210) × 7674 for GOSAT (in situ + GOSAT)) to be successfully inverted on our system (with a single CPU).

3. Results

3.1. General evaluation of inversions, including short-term flux variability

Posterior model mole fractions are closer to the assimilated observations than are the prior mole fractions for the in situ-only, GOSAT-only, and in situ + GOSAT inversions, as desired, as suggested by Fig. 2 and indicated by the means and standard deviations of the model-observation differences over all observations shown in Fig. 3 (a, d, e, and f). Comparison of

posterior mole fractions with the data set not used (Fig. 3b, c), on the other hand, gives mean differences not as close to 0 as in the comparison with the assimilated data (Fig. 3d and 3a, respectively), and standard deviations that are larger than for the prior; this reflects the fact that the in situ and GOSAT data sets are not necessarily consistent with each other and combine to produce larger standard deviations than with the less variable prior model, which has not assimilated any data~~have independent random errors~~. The improved agreement between model and assimilated observations is reflected also in the ~~chi-squared (~~cost function~~)~~ values before and after the inversions shown in Table 2. The ~~expected~~ minimized cost function ~~value of~~ follows a χ^2 ~~chi-squared~~ distribution~~(normalized),~~ and should thus have a value close to 1 (normalized by the number of observations) for a satisfactory inversion ~~is 1, which signifies that the tightness of the fit of the results to the observations and to the a priori parameter estimates is comparable to the level of uncertainty assumed for the observations and the a priori estimates~~(Tarantola, 1987; Rayner et al., 1999). The posterior ~~chi-squared~~ cost function values for all of the inversions are closer to 1 than the prior values.

In addition to cross-evaluating the in situ-only and GOSAT-only inversions, we evaluate both inversions against the independent, well-calibrated Amazon aircraft data set, which samples an under-observed region with large, variable fluxes. Vertical profiles of the model and the aircraft data (Fig. S1 in the supplementary material) show that the prior mole fractions often exhibit a bias relative to the aircraft observations, especially in a boundary layer-like structure below ~2 km altitude, with the sign of the average bias varying from season to season. The in situ inversion ~~often~~ exhibits worse agreement with the observations than the prior does more often than it is better (e.g. with a root mean square error (RMSE) that is more than 1 ppm larger in 27 of 60 cases above 2 km~~,~~ and in 27 cases below 2 km, and more than 1 ppm smaller in only

12 cases above 2 km and 14 cases below 2 km). The GOSAT inversion ~~not only~~ exhibits smaller discrepancies with the observations than the in situ inversion does ~~in general~~ more often than the reverse, in both altitude ranges~~(29 of 60 cases above 2 km and 28 cases below 2 km, while the reverse is true in only 11 cases above 2 km and 25 cases below 2 km),~~. ~~but~~ Furthermore, ~~it~~ the GOSAT inversion ~~also~~ is more often better than the prior than worse above 2 km ~~(16 vs. 13 cases)~~. ~~In contrast, the in situ inversion is more often worse than the prior than better (e.g. 27 vs. 12 cases above 2 km).~~ Overall statistics, computed separately for lower and higher altitudes, are shown in Fig. 4. The model-observation histograms indicate that agreement with the aircraft observations is again better for the GOSAT inversion than the in situ inversion, with smaller or comparable mean differences and standard deviations. There is a near complete lack of in situ sites in the inversion that are sensitive to Amazon fluxes (as suggested by Fig. 1a), contrasting with the availability of some GOSAT data over the region (Fig. 1b), meaning that regional flux adjustments in the in situ inversion are driven, often erroneously, by correlations with fluxes outside of the region (as will be discussed in depth below in Sect. 3.3). The GOSAT inversion agrees with the aircraft observations better than the prior does above 2 km, implying that incorporating GOSAT data in the inversion results in better performance than no data. However, the posterior model-observation differences have greater variance than the prior below 2 km. A possible explanation for this is that the use of GOSAT observations in an inversion introduces more random error in the model mole fractions; given that the GOSAT data are sparse over the Amazon, there is little data averaging over the 8-day intervals and flux regions and random errors can thus have a substantial impact~~, which would not be expected to be correlated with the aircraft measurement error~~. GOSAT errors presumably affect higher altitudes in the model less,

since the mole fractions there are influenced by fluxes across a broader area than at lower altitudes and thus errors are averaged out to a greater extent.

Example time series of 8-day mean prior and posterior NEP and ocean fluxes for the in situ-only and GOSAT-only inversions are shown in Fig. 5. Since the posterior fluxes in our inversion regions tend to have large fractional (percentage) uncertainties, especially for the in situ-only inversion, we focus in this paper on results aggregated to larger regions. To facilitate comparison with other studies, results are aggregated to TC3 land and ocean regions, accounting for error correlations. The posterior time series exhibit larger fluctuations than the prior time series, especially for the in situ inversion over land. The fluctuations would presumably be smaller if we excluded flagged, outlier in situ observations or used a smoothed data product such as GLOBALVIEW-CO_2 (2009), which has been used in many inversions including those of TC3 and some of those in the Houweling et al. (2015) intercomparison. In addition, some of the fluctuations likely represent actual variability in the fluxes, while other fluctuations are probably noise. In fact, the calculated numbers of degrees of freedom for signal and noise (as defined by Rodgers, 2000) are 3525 and 4186 for the in situ inversion (summing up approximately to the number of inversion parameters, 7674) and 4925 and 2947 for the GOSAT inversion. This indicates that ~45% of the in situ inversion solution is based on actual information from the measurements, given the assumed prior and observation uncertainties, while ~65% of the GOSAT inversion solution is constrained by the measurements. The in situ data set is sparser than GOSAT, especially over land, and thus contains greater spatial sampling bias, so that many of the flux regions are under-determined and may exhibit so-called dipole behavior associated with negative error correlations (discussed further below).

Results for the in situ + GOSAT inversion (not shown in Fig. 5) lie mostly in between the in situ-only and GOSAT-only results. The fluxes generally lie closer to those of the GOSAT-only inversion for regions with a relatively low density of in situ measurements, including tropical and southern land regions, while they lie closer to those of the in situ-only inversion for regions with a relatively high density of in situ measurements, including northern land and many ocean regions. As expected, there are a larger number of degrees of freedom for signal, 6553, than for either the in situ-only or the GOSAT-only inversion (and fewer degrees of freedom for noise, 1632), indicating that the two data sets provide a certain amount of complementary information. Here, ~80% of the inversion solution is constrained by the measurements.

To average out noise in the posterior fluxes and to better observe the major features in the results, we show monthly average fluxes in Fig. 6. There is a similar onset of seasonal CO_2 drawdown in the GOSAT-only inversion and the CASA-GFED prior in Boreal North America, Temperate North America, and Boreal Asia, whereas the in situ-only inversion is noisier, similar to what was noted above. The GOSAT inversion suggests an overall shift in the global CO_2 sink from tropical and southern land to northern land regions relative to the prior and the in situ inversion, similar to what has been found in previous GOSAT inversions (e.g. Houweling et al., 2015; discussed further below). There are some unusual features in the GOSAT inversion. For example, there is a negative flux in January in some northern regions, with the 1σ range lying entirely below zero for Boreal Asia and Europe; this CO_2 uptake does not seem plausible in the middle of winter for these regions. Also, there are large positive fluxes during winter through spring in Northern Africa, which deviate from the prior beyond any overlap in the 1σ ranges for two months and whose 1σ ranges stay above zero for six months, summing up to a source of 1.9 Pg C over the period December through May, not including fires. The fluxes are larger than

514 those of any sustained period of positive fluxes in any region in either the prior or the in situ
515 inversion. The anomalous features suggest that the GOSAT inversion is affected by uncorrected
516 retrieval biases that vary by season and region (as has been shown by Lindqvist et al. (2015) and
517 Kulawik et al. (2016)) and/or sampling biases, including a lack of observations at high latitudes
518 during winter, which limit the ability to accurately resolve inferred fluxes down to the scale of
519 TransCom regions.

520 A comparison of the monthly mean fluxes with those from another inversion system,
521 NOAA's CarbonTracker version 2013B (CT2013B), is displayed in Fig. 7. CT2013B is an
522 ensemble Kalman smoother data assimilation system with a window length of five weeks that
523 uses multiple in situ observation networks and prior models to optimize weekly fluxes (Peters et
524 al., 2007; https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2013B/CT2013B_doc.php,
525 accessed 4 October 2016). Similar to the present study, CT2013B uses CASA-GFED3 fluxes
526 from van der Werf et al. (2010) as one of the land NEP priors, though with different FPAR and
527 meteorological driver data. (CASA-GFED2 is the other land prior in its ensemble of priors.) In
528 addition, CT2013B uses the seawater pCO_2 distribution from the Takahashi et al. (2009)
529 climatology to compute fluxes for one of its ocean priors; the other ocean prior is based on
530 results from an atmosphere-ocean inversion. CT2013B uses 93 observation time series while 87
531 are used here, although the former include measurements by multiple labs at the same site and
532 flask and continuous measurements at the same site (where duplicate observations are de-
533 weighted by inflating the model-data mismatch error by the square root of N). A notable
534 difference is that CT2013B solves for uniform flux scale factors over entire ecosystem types
535 within each TC3 region, with the ecosystem types not necessarily being contiguous. Results
536 from our in situ-only inversion are shown alongside those of CT2013B in Fig. 7 aggregated over

large regions. The two sets of posterior flux time series are similar overall, with overlapping 2σ ranges at all times except in the extratropical northern oceans~~Northern Oceans~~ region. One distinctive feature is that the posterior fluxes stay closer to the priors for CT2013B. A likely explanation is the tighter prior uncertainties in CT2013B, the magnitudes of which are on average 40% of ours for land regions and 30% of ours for ocean regions. For its ocean prior based on an atmosphere-ocean inversion, CT2013B assumes uncertainties consistent with the formal *posterior* uncertainties from the inversion, which are relatively small because of the large number of ocean observations used in the inversion; uniform fractional uncertainties are assumed for the other ocean prior and the land priors. Another feature is the larger month-to-month fluctuations in our results. ~~A number of factors could contribute to smaller fluctuations in the CT2013B results.~~ In addition to~~One of them is~~ the tighter prior uncertainties used,~~.~~ a~~A~~nother factor that could contribute to smaller fluctuations in the CT2013B results is the use of prior estimates that represent a smoothing over three assimilation time steps, which attenuates variations in the forecast of the flux parameters in time. And another factor is that to dampen spurious noise due to the approximation of the covariance matrix by a limited ensemble, CT2013B applies localization for observation sites outside of the marine boundary layer, in which flux parameters that have a non-significant relationship with a particular observation are excluded. We further evaluate our inversions in the following sections.

3.2. Longer-term budgets and observation biases

Longer-timescale budgets can be assessed in Fig. 8, which displays 12-month mean fluxes (Jun 2009-May 2010) over large, aggregated regions, with fires now included, for our inversions and the CT2013B inversion. Results for individual TC3 regions are shown in Table 1.

The global total flux (including fossil emissions) is substantially more positive for the GOSAT-only inversion relative to the in situ-only inversion, 6.5 ± 0.2 Pg C y^{-1} vs. 4.1 ± 0.5 Pg C y^{-1}, while that for the in situ + GOSAT inversion lies in between at 5.7 ± 0.2 Pg C y^{-1}. Such a large difference in the atmospheric CO_2 growth rate implied by the two distinct data sets is plausible even if there are no trends in uncorrected biases between the data sets, given their sampling of different regions of the atmosphere and the relatively short 12-month time frame over which the growth occurs. In fact, for a different 12-month period within our analysis, Sep 2009-Aug 2010, the total fluxes for the GOSAT-only and in situ-only inversions are much closer to each other– 5.53 Pg C y^{-1} and 5.47 Pg C y^{-1}. Houweling et al. (2015) also found a larger total flux in the GOSAT-only inversions relative to the in situ during Jun 2009-May 2010 averaged across 8 models~~,~~, ~4.8 Pg C y^{-1} vs. ~4.6 Pg C y^{-1}, with a substantial amount of inter-model variability within those averages~~though the difference may not be statistically significant~~.

As was noted ~~earlier~~ in ~~this~~ ~~S~~section 3.1, the GOSAT-only inversion exhibits a shift in the global CO_2 sink from tropical and southern land to northern land relative to the prior and the in situ-only inversion (Fig. 8). The differences are within the 1σ uncertainty ranges. The shift includes an increase in the source in N. Africa (0.2/1.5/2.0 Pg C y^{-1} for prior/in situ-only/GOSAT-only), Temperate S. America (0.4/0.4/1.1 Pg C y^{-1}), and Australia (0.0/-0.2/0.6 Pg C y^{-1}), and an increase in the sink in Europe (-0.1/0.6/-1.5 Pg C y^{-1}) and Temperate N. America (-0.3/-0.6/-1.5 Pg C y^{-1}) (Table 1). As for the ocean, the GOSAT inversion also exhibits a larger source in the tropics relative to the prior and the in situ inversion (outside of the 1σ ranges; Fig. 8). However, the GOSAT inversion now exhibits a smaller sink over extratropical northern oceans relative to the in situ inversion, and a larger sink over extratropical southern oceans relative to both the prior and the in situ inversion (at or outside of the 1σ ranges). The TC3

583 regions contributing the most to these differences include Tropical Indian (0.1/0.0/0.7 Pg C y^{-1}
584 for prior/in situ-only/GOSAT-only), N. Pacific (-0.9/-0.5 Pg C y^{-1} for in situ-only/GOSAT-only),
585 N. Atlantic (-0.8/-0.5 Pg C y^{-1} for in situ-only/GOSAT-only), and Southern Ocean (-0.2/-0.4/-0.9
586 Pg C y^{-1} for prior/in situ-only/GOSAT-only) (Table 1).

587 The GOSAT results appear to contradict global carbon cycle studies that favor a weaker
588 terrestrial net source in the tropics compensated by a weaker northern extratropical sink (e.g.
589 Stephens et al., 2007; Schimel et al., 2015). We show the north-south land carbon flux
590 partitioning of our results in Fig. S2 in the manner of Schimel et al. (2015). The shift in the sink
591 from the south + tropics to the north in the GOSAT inversion relative to the in situ inversion
592 goes in a direction opposite to that consistent with an airborne constraint considered by Stephens
593 et al. (2007) and with the expected effect of CO_2 fertilization according to Schimel et al. (2015).
594 However, the shift may be due at least in part to GOSAT retrieval and sampling biases. An
595 evaluation of posterior mole fractions in the GOSAT-only inversion against surface in situ
596 observations indicates that the GOSAT inversion may be biased low during much of the analysis
597 period over Europe and Temperate N. America, especially in winter (when there is little direct
598 constraint at high latitudes by GOSAT observations), and biased somewhat high over N. Africa,
599 especially in spring. However, the dearth of in situ sites over N. Africa, with only one in the
600 middle of the region (~~Assekrem,~~in Algeria) and a few around the edges (e.g. ~~Izaña,~~ Canary
601 Islands and ~~Mt. Kenya,~~ Kenya), precludes a definitive evaluation over that region. Globally, the
602 GOSAT inversion tends to underestimate mole fractions at high latitudes of the Northern
603 Hemisphere, often by more than 1σ, as shown by latitudinal profiles averaged over all surface
604 sites by season (Fig. 9), suggesting an overestimated northern sink. The same is true of the high
605 latitudes of the Southern Hemisphere. The GOSAT inversion overestimates mole fractions in

parts of the tropics, sometimes by more than 1σ (Fig. 9), suggesting an overestimated tropical source. Uncorrected retrieval biases may be especially prevalent in the tropics, where there are very few TCCON stations available as input to the GOSAT bias correction formulas; only 1 TCCON station, Darwin, Australia, was operating in the tropics during 2009-2010, and only 2 more stations, Reunion Island and Ascension Island, became operational during the rest of the ACOS B3.4 retrieval period. In contrast, the posterior mole fractions for the in situ-only inversion generally agree well with the surface observations (Fig. 9; also seen in the individual site time series in Fig. 2), which is expected given that these are the observations that are used in the optimization. The prior mole fractions are generally too high, which is consistent with the fact that the CASA-GFED biosphere is near neutral while the actual terrestrial biosphere is thought to generally be a net CO_2 sink.

Evaluation of the inversions against latitudinal profiles constructed from HIPPO aircraft measurements, which provide additional sampling over the Pacific, does not indicate any widespread overestimate by the GOSAT inversion relative to the observations in the tropics (Fig. 10a-f), unlike what was seen in the comparison with the more globally distributed surface observations~~indicates an overestimate by the GOSAT inversion relative to HIPPO in parts of the tropics at lower altitudes (Fig. 10a, d), similar to what was seen in the comparison with surface observations~~. ~~And~~ But the GOSAT inversion does exhibit~~s~~ an underestimate relative to HIPPO ~~in the southern extratropics~~from ~40°S southward in the lower to middle levels of the troposphere (Fig. 10a, b, d, e), especially for Mission 2 (Oct-Nov 2009). Again, retrieval bias and sampling bias (~~the~~ a lack of GOSAT ocean observations south of ~40°S and land observations south of ~50°S~~at southern high latitudes~~) ~~points to sampling bias as the~~are likely the cause~~s~~ of the underestimate. In the northern extratropics, ~~the comparison of the inversions with~~

~~HIPPO has different features from the comparison with surface data:~~ the GOSAT inversion actually ~~generally~~ exhibits higher mean mixing ratios than HIPPO in general in the lower troposphere, especially for Mission 2, and the in situ inversion gives higher mixing ratios than HIPPO at some latitudes ~~in the northern extratropics~~ and lower mixing ratios at others for Mission 2. In one particular latitude range, 55-67°N, both inversions give much higher mixing ratios than HIPPO, by up to 67 ppm in the case of the in situ inversion and 30 ppm for the GOSAT inversion. This could reflect inaccuracy in posterior fluxes due to the inversions' being under-constrained over the high-latitude North Pacific and Alaska, with few observations during this season in the case of GOSAT and a tendency for the sparse in situ network to produce noisy inversion results, as was discussed above. However, given that the prior model also gives substantially higher mixing ratios than HIPPO at these latitudes (by up to 11 ppm), the discrepancy could be due in part to some factor common to the prior and posteriors such as model transport or representation error.

In the upper troposphere to lower stratosphere, the GOSAT inversion more often than not exhibits better agreement with the HIPPO observations than the in situ inversion does for both Mission 2 and 3 (Fig. 10c, f). (We think it is reasonable to include data from these altitudes as part of the evaluation of the inversion results, since the tropopause in the GEOS-5/MERRA meteorological data assimilation system underlying PCTM transport is considered to be accurate (Wargan et al., 2015) and PCTM has been shown to simulate upper troposphere-lower stratosphere trace gas gradients well compared to other models (Patra et al., 2011).) A likely explanation is that the GOSAT data provide constraints throughout the atmospheric column, whereas the in situ measurements constrain only surface CO_2. Figure 10 shows that the high-altitude mole fractions from the in situ inversion are consistently close to those of the prior,

suggesting that the lack of high-altitude constraints prevents major adjustments in mole fractions at these levels, unlike in the GOSAT inversion. ~~Although~~ Regarding potential impacts of GOSAT biases ~~may affect these altitudes as well~~, ~~it appears to be the case that the data still provide better constraints than no observations above the surface at all. Furthermore,~~ an air parcel at higher altitudes, especially in the upper troposphere or above, generally consists of a mixture of air originating from a broad area near the surface (e.g. Orbe et al., 2013), and thus the effects of different regional biases in posterior fluxes may cancel out at those altitudes.

The conclusion that GOSAT biases may contribute to the shift in the land sink is also supported by Houweling et al. (2015). That study reported a shift in the GOSAT-only inversions relative to the in situ inversions consisting of an increase in the sink in northern extratropical land of 1.0 Pg C y^{-1} averaged across models and an increase in the source in tropical land of 1.2 Pg C y^{-1} during June 2009-May 2010; in comparison, our inversions produce an increase in the northern land sink of 0.4 Pg C y^{-1} and an increase in the tropical land source of 1.2 Pg C y^{-1} (Fig. 8). Houweling et al. (2015) found an especially large and systematic shift in flux of ~0.8 Pg C y^{-1} between N. Africa and Europe, but then provided evidence that the associated latitudinal gradient in CO_2 mole fractions may be inconsistent with that based on surface and HIPPO aircraft in situ observations. They also suggested that the shift in annual flux between the two regions may be a consequence of sampling bias, with a lack of GOSAT observations at high latitudes during winter. Chevallier et al. (2014) also found a large source in N. Africa of ~1 Pg C y^{-1} in their ensemble of GOSAT inversions and considered the magnitude of that unrealistic, given that emissions from fires in that region likely amount to < 0.7 Pg C y^{-1}. (Note that our N. Africa source is even larger than that of Chevallier et al. (2014).) Inversion experiments by Feng et al. (2016) provide evidence that the large European sink inferred from GOSAT observations

may be an artifact of high XCO_2 biases outside of the region that necessitate extra removal of CO_2 from incoming air for mass balance, in concert with sub-ppm low XCO_2 biases inside the region. An observing system simulation experiment by Liu et al. (2014) found that GOSAT seasonal and diurnal sampling biases alone could result in an overestimated annual sink in northern high-latitude land regions. And a review paper by Reuter et al. (2017) further highlighted the discrepancy between satellite-based and ground-based estimates of European CO_2 uptake and cited retrieval and sampling biases as possible sources of error in the former (while also noting sampling issues with in situ networks for the region).

Again, the results for the in situ + GOSAT inversion lie mostly in between those for the in situ-only and GOSAT-only inversions, with the in situ + GOSAT fluxes lying closer to the GOSAT-only ones for the tropical/southern land regions and land as a whole (Table 1 and Fig. 8), suggesting the dominance of the GOSAT constraint in these regions. The posterior uncertainties for the GOSAT inversion (Table 1) are as small as or smaller than those for the in situ inversion, except in Boreal and Temperate N. America, N. Pacific, Northern Ocean, and Southern Ocean. This reflects the fact that GOSAT generally provides better spatial coverage, except over N. America, where the in situ network provides good coverage, and over and near high-latitude ocean areas, where there is decent in situ coverage and poor GOSAT coverage. Uncertainty reductions in the in situ inversion range from 15% to 93% for land regions and 15% to 56% for ocean regions (Table 1). In the GOSAT inversion, the uncertainty reductions range from 43% to 89% for land and 19% to 56% for ocean. And in the inversion with combined in situ and GOSAT data, the uncertainty reductions are larger than or equal to those in either the in situ-only or the GOSAT-only inversion, ranging from 61% to 96% for land and 40% to 67% for ocean.

3.3. Flux error correlations and land-ocean partitioning

Here we elaborate on ~~the subject of~~ posterior error correlations, which indicate the degree to which fluxes are estimated independently of one another. Negative correlations can be manifested in dipole behavior, in which unusually large flux adjustments of opposite signs occur in neighboring regions/time intervals. These are shown in Fig. 11 aggregated to TC3 regions and the 12-month period from June 2009 to May 2010. The full-rank error covariance matrix generated by the exact Bayesian inversion method (from which the correlation coefficients are derived) is a unique product of this study, particularly as applied to satellite data. There are a larger number of sizable correlations between land regions in the in situ inversion than in the GOSAT inversion (in the top left quadrants of the plots). One specific feature is negative correlations among the four TC3 regions in South America and Africa ("Trop Am", "Temp S Am", "N Africa", and "S Africa") in the in situ inversion, whereas in the GOSAT inversion there are negative correlations within South America and within Africa but not between the two continents. Although there are less extensive correlations over land in the GOSAT inversion, they are often of larger magnitude than in the in situ inversion; this could reflect the fact that GOSAT observations, though of higher density than the in situ observations over many regions, are column averages representing mixtures of air from a broader source region than for surface observations, and may thus result in larger error correlations for immediately adjacent regions, e.g. within a continent. Over the ocean regions, ~~In~~ in contrast, the GOSAT inversion exhibits anti-correlations that are as extensive as those for the in situ inversion and often of~~exhibits~~ larger magnitude~~anti-correlations over the ocean regions~~. For example, there are substantial negative correlations between Southern Ocean and each of the other southern regions—S. Pacific, S.

721 Atlantic, and S. Indian. This is consistent with the almost complete lack of GOSAT observations
722 at the latitudes of the Southern Ocean region and the southern edges of the neighboring ocean
723 regions (Fig. 1b). Interestingly, there is not a sizable correlation between N. Africa and Europe
724 in the GOSAT inversion (in either seasonal or 12-month means), which runs counter to what
725 might be expected from the shift in flux discussed above; rather, each of these regions is
726 correlated with a number of other regions. We do find a fairly large correlation of -0.62 between
727 the northern extratropics in aggregate (land + ocean) and the tropics for the 12-month period
728 though. Correlations for the in situ + GOSAT inversion (not shown) generally lie in between
729 those of the in situ-only and GOSAT-only inversions. Even with the incorporation of both sets
730 of observations, there are substantial correlations of as much as -0.6 between regions within a
731 continent, reinforcing our earlier conclusion that sampling gaps limit the ability of the
732 observations to constrain fluxes down to the scale of most TC3 regions.

733 The in situ-only and CT2013B posterior global totals are nearly the same, but the land-
734 ocean split is different, with our inversion exhibiting a larger sink over ocean than over land
735 (with non-overlapping 2σ ranges) while in CT2013B the land and ocean fluxes are similar, with
736 the ocean flux changing little from the prior (Fig. 8). A likely explanation for the difference is
737 the very tight prior constraints on ocean fluxes of CT2013B that were discussed above, which
738 force the flux adjustments to take place mostly on land. The GOSAT inversion also exhibits a
739 relatively large ocean sink of -3.1 ± 0.5 Pg C y^{-1}; for comparison, the CT2013B estimate is -2.4 ±
740 0.4 Pg C y^{-1}, our in situ-only estimate is -4.0 ± 0.8 Pg C y^{-1}, and the estimate of the Global
741 Carbon Project (GCP) is -2.5 ± 0.5 Pg C y^{-1} for 2009-2010 (Le Quéré et al., 2013; Le Quéré et
742 al., 2015). The GCP estimate is a synthesis that combines indirect observation-based estimates
743 for the mean over the 1990s with interannual variability from a set of ocean models and accounts

744 for additional observation-based estimates in the uncertainty. The difference between our
745 inversion estimates and the GCP estimate is actually even larger than suggested by those
746 numbers, given that a background river to ocean flux of ~0.5 Pg C y^{-1} should be subtracted from
747 our ocean flux to make it comparable to the GCP ocean sink, which refers to net uptake of
748 *anthropogenic* CO_2 (Le Quéré et al., 2015). Our relatively small land sink is reflected in our
749 inversion results' lying mostly outside of the GCP global land flux range in the north-south
750 partitioning plot in Fig. S2. Similarly, in comparing our results with those of Houweling et al.
751 (2015), we find that the global budgets are comparable for all three inversions—in situ-only,
752 GOSAT-only, and in situ + GOSAT—as was mentioned above, but the land-ocean split is
753 different. Our posterior ocean flux is -4.0 ± 0.8 Pg C y^{-1}, -3.1 ± 0.5 Pg C y^{-1}, and -3.9 ± 0.3 Pg C
754 y^{-1} for the three inversions, while it is -1.6 ± 0.5 Pg C y^{-1}, -1.2 ± 0.6 Pg C y^{-1}, and -1.5 ± 0.8 Pg C
755 y^{-1} in the results of Houweling et al. (2015; pers. comm., 2016) (averaged over different
756 weighted averages of the models).

757 There is a strong negative correlation globally between posterior flux errors for land and
758 ocean of -0.84 and -0.89 in the in situ-only and the GOSAT-only inversion, respectively. Basu et
759 al. (2013) also reported a large negative correlation between land and ocean fluxes of -0.97 in
760 their in situ + GOSAT inversion during September 2009-August 2010. The anti-correlations
761 imply that the observations cannot adequately distinguish between adjustments in the global land
762 and ocean sinks. Thus, land-ocean error correlation may be a fundamental challenge that global
763 CO_2 flux inversions are faced with, at least given the sampling characteristics of the in situ and
764 GOSAT data sets used here. Without tight prior constraints on ocean fluxes, those fluxes are
765 subject to large, and potentially unrealistic, adjustments (i.e. dipole behavior).

766 To assess the effect of prior constraints on the inversion, we conducted a test with
767 reduced prior uncertainties, for both land and ocean fluxes, so that they are similar on average to
768 those of CT2013B. Results for an in situ-only inversion and a GOSAT-only inversion are shown
769 in Table 1 and Fig. 12. For the in situ-only inversion, the posterior ocean flux is now much
770 smaller in magnitude, -2.8 ± 0.3 Pg C y^{-1}. The posterior ocean flux for the GOSAT inversion
771 does not change as much, decreasing in magnitude from the original -3.1 ± 0.5 Pg C y^{-1} to -2.9 ±
772 0.2 Pg C y^{-1}. The ocean flux 1σ ranges for both inversions now overlap with the 1σ range ~~that~~ of
773 CT2013B; accounting for the riverine flux, the 1σ range for the in situ inversion overlaps with
774 the 1σ range ~~that~~ of GCP, while the 1σ range for the GOSAT inversion is still just outside of that
775 of GCP. The better agreement with the GCP budget (land component) can also be seen in Fig.
776 S2 for both inversions. The inversions with tighter priors have slightly larger cost function
777 values than the baseline inversions (Table 2; the difference for the GOSAT cases is concealed by
778 rounding). The inversions with tighter priors generally exhibit better agreement with
779 independent observations, e.g. lower-altitude HIPPO observations (Fig. ~~13~~S3), and surface
780 observations in the case of the GOSAT inversion (Fig. ~~14~~S4), indicating that the effects of
781 sampling and retrieval biases are reduced with tighter prior uncertainties. The better agreement
782 also lends support to the smaller ocean sink estimates. (At high altitudes, keeping posterior mole
783 fractions closer to the prior mole fractions results in worse agreement with HIPPO in many
784 places, especially for the GOSAT inversion.) However, the tighter priors do not completely
785 eliminate the discrepancies between the inversions and the independent observations, suggesting
786 that tight priors may not completely counteract the effects of observational biases.
787 Basu et al. (2013) saw a similar underestimate of mole fractions during parts of the year
788 in the southern extratropics in their GOSAT inversion relative to surface observations and

overestimate of the seasonal cycle, though with some differences in the shape of the seasonal cycle from our study (including a later descent toward and recovery from the annual minimum in austral summer and a larger peak in late winter-early spring). They, however, used the SRON-KIT RemoTeC GOSAT retrieval with a known issue over the ocean, and concluded that adding global land and ocean observation bias correction terms to their inversion was needed to make the land-ocean flux split more realistic and to improve the seasonal cycle of CO_2 in the southern extratropics. In contrast, studies have found no noticeable bias in the ACOS B3.5 ocean glint XCO_2 retrievals relative to TCCON (Kulawik et al., 2016) and a mean bias of only -0.06 ppm relative to HIPPO (Frankenberg et al., 2016); the B3.4 version we use is on average ~0.2 ppm lower than B3.5 in 2010 (Deng et al., 2016). So although a small overall negative bias in the bias-corrected ACOS B3.4 ocean data cannot be ruled out (and there could of course be larger negative biases on a regional scale, such as in the southern extratropics), we conclude that the land-ocean flux split in inversions using either in situ or GOSAT data is strongly influenced by error correlations and dependent on the prior uncertainties assumed.

The shift in the global terrestrial sink from the tropics/south to the north when comparing the GOSAT-only inversion with the in situ-only inversion and the prior is still seen when prior uncertainties are decreased (Fig. 12; Fig. S2), as is a substantially more positive~~larger~~ global total budget in the GOSAT inversion relative to the in situ (Fig. 12). The uncertainty reductions in the test inversions are smaller than those in the baseline inversions (Table 1), as is expected from the smaller starting values of the uncertainties. In summary, the magnitude of the ocean sink and the partitioning of the global sink between land and ocean are sensitive to the prior uncertainties, but other inferred features of the carbon budget are robust with respect to prior uncertainties.

812 Given that there is uncertainty in the land-ocean flux partitioning at sub-global scales as
813 well (e.g. as indicated by moderate negative correlations between northern land and northern
814 oceans, tropical land and tropical oceans, etc.), we consider results for combined land and ocean
815 regions in Figs. 8 and 12. They indicate that there is a shift in the global sink from the tropics to
816 the north and the south in the GOSAT inversion relative to the prior, and an increased source in
817 the tropics of ~2 Pg C y^{-1} in the GOSAT inversion relative to the in situ inversion. These
818 features are seen in the inversions with tighter priors as well as in the baseline inversions. Note
819 that the increased source over southern land and increased sink over southern oceans~~southern~~
820 ~~ocean~~ in the GOSAT inversion relative to the in situ inversion that were discussed earlier cancel
821 each other out approximately, suggesting a compensation of errors. Also note that the inversion
822 using the in situ + GOSAT data sets, which provide more constraint than either of the data sets
823 alone, produces a global flux close to mid-way between the in situ-only and GOSAT-only
824 inversions, while it produces a Tropic Land + Oceans flux much closer to that of the GOSAT
825 inversion than to the in situ inversion. This suggests some degree of independence of the
826 GOSAT-inferred regional result from the global result.
827
828 3.4. Impacts of climatic conditions on 2009-2010 fluxes
829 We now analyze the impacts of several climatic events during the analysis period on CO_2
830 fluxes as indicated by the inversion results. We focus on 1) unusually hot and dry conditions at
831 Northern Hemisphere higher latitudes in summer of 2010, 2) wetter conditions over parts of
832 North America in spring and early summer of 2010 relative to 2009, and 3) record drought in the
833 Amazon in 2010.

Guerlet et al. (2013), who examined GOSAT data and performed a flux inversion using a variational assimilation system, found that there was less net terrestrial CO_2 uptake in summer of 2010 than in 2009 at northern high latitudes, consistent with known severe heat waves, drought, and high fire emissions, especially across Eurasia, centered around western Russia, and to a lesser extent in North America.

Motivated by that study, we examined our inversion results for 2009 and 2010, focusing on the GOSAT inversion. As can be seen in the global maps of natural plus biomass burning fluxes in June-July-August (JJA) in Fig. 1~~3~~5, the GOSAT inversion does appear to exhibit a decreased CO_2 uptake over Eurasia, including the area around western Russia (enclosed in a box in the figure), in 2010. A decreased sink can also be seen in parts of North America. A decreased sink over western Russia can also be seen in the CASA-GFED prior, though of a smaller magnitude. In contrast, there is actually an increased sink in that region in the in situ inversion. In fact, none of the sites used are in or immediately downwind of that region (Fig. 1a). Total NEP and fire fluxes over northern TC3 regions are shown in Fig. 1~~4~~6. There is less CO_2 uptake in JJA 2010 than in 2009 in all the regions except Temperate Asia in the GOSAT-only inversion. The differences exceed the 1σ ranges for 3 of the 5 regions, even exceeding the 3σ ranges for Europe, which includes western Russia. Also shown is the in situ + GOSAT inversion, which exhibits a similar pattern of 2010-2009 differences. These inversion results are thus consistent with the earlier GOSAT study. In contrast, the 2010-2009 differences in the prior are small and, for some regions, of the opposite sign as that in the inversions (Fig. 1~~4~~6).

Measurements from the JR-STATION tower network are suitably located for evaluating the inferred flux interannual variability over Eurasia. Time series are shown in Fig. 1~~5~~7 for observations, the prior model, and the GOSAT-only inversion at 6 sites with complete

857 summertime data in 2009-2010. (As with the continuous measurements used in the in situ
858 inversion, afternoon data are selected to avoid difficulties associated with nighttime boundary
859 layers.) Posterior mole fractions are noisier in the wintertime, likely a result of the lack of
860 GOSAT observations during that season at these high latitudes. Focusing on 2010-2009
861 differences, the observations suggest a shallower drawdown in 2010 than in 2009 at most of the
862 sites, which is generally captured by both the prior and the GOSAT posterior. It appears though
863 that the GOSAT inversion exaggerates the 2010-2009 difference at some of the sites,
864 overestimating especially the drawdown in 2009. For a more quantitative analysis, we calculate
865 the average 2010-2009 difference in mole fractions over June-July-August for each site (Table
866 3). The GOSAT-only inversion overestimates the 2010-2009 difference at 5 of the 6 sites. The
867 in situ + GOSAT inversion exhibits less of an overestimate overall than the GOSAT-only
868 inversion, with 3 of the 6 sites being substantially overestimated. The prior exhibits the best
869 agreement with the observations overall.

870 The earlier study by Guerlet et al. (2013) assumed that the differences between 2010 and
871 2009 posterior biospheric fluxes are relatively insensitive to biases in the GOSAT data, since at
872 least some of those errors may be similar between the two years. However, our evaluation of the
873 inversions using JR-STATION data suggests that retrieval biases can vary significantly from
874 year to year. Kulawik et al. (2016) estimated a year-to-year variability in GOSAT biases relative
875 to TCCON of 0.3 ppm averaged over the stations. Another study has raised a separate but
876 related issue of inversion results potentially being sensitive to the spatiotemporal distribution of
877 observations in different data sets (e.g. different GOSAT retrievals) (H. Takagi, pers. comm.,
878 2015); by extension, comparison of fluxes from two time periods can be affected by changes in
879 the distribution of observations over time within a particular data set. But in JJA 2009 and 2010,

880 there are similar numbers of ACOS GOSAT observations overall in the northern land region, so
881 differences in data coverage are probably not a factor in this particular case study.
882 Our evaluation using JR-STATION data also indicates that the prior may be a reasonable
883 estimate of the 2010-2009 difference in growing season fluxes, at least over Siberia, despite
884 possible shortcomings in the simulation of drought impacts on NEP and of the overall magnitude
885 of fire emissions by CASA-GFED3. The latest version of GFED (version 4s), which includes
886 small fires, tends to generate higher emissions than GFED3 (van der Werf et al., 2017).
887 Over large parts of North America, conditions were wetter in spring and early summer of
888 2010 than in 2009, especially in the western half of the U.S. and adjacent parts of Mexico and
889 Canada, as suggested by North American drought maps for June 2010 vs. June 2009 (e.g.
890 https://www.drought.gov/nadm/content/map/2010/06) and shallow groundwater status maps for
891 the U.S. based on GRACE (Gravity Recovery and Climate Experiment) satellite data for May-
892 June (Houborg et al., 2012; http://droughtcenter.unl.edu/NASA/GRACE/). Consistent with the
893 wetter conditions in 2010 are a larger CO_2 sink over North America (Boreal + Temperate) in
894 May-June 2010 relative to 2009 in our priors (-5.0 ± 3.9 Pg C y^{-1} vs. -3.4 ± 3.9 Pg C y^{-1}), in situ-
895 only posteriors (-5.0 ± 0.4 Pg C y^{-1} vs. -3.8 ± 0.5 Pg C y^{-1}), and GOSAT-only posteriors (-5.8 ±
896 0.4 Pg C y^{-1} vs. -3.3 ± 1.8 Pg C y^{-1}). We consider the in situ inversion result to be reliable here,
897 given the large uncertainty reduction for North America and small error correlations with other
898 regions (not shown). The 2010 and 2009 fluxes differ such that their 1σ ranges do not overlap
899 for the in situ and the GOSAT posteriors. Much warmer conditions in eastern North America in
900 May-June 2010 compared to 2009 (e.g. https://www.ncdc.noaa.gov/sotc/global/201005 and
901 https://www.ncdc.noaa.gov/sotc/global/200905) may have also contributed to increased uptake,
902 especially at higher latitudes, where insufficient warmth can be more of a limiting factor for NEP

than insufficient moisture during late spring-early summer. Despite the increased sink in June 2010 over North America, the 2010 summer exhibits a decreased sink relative to 2009 when integrated through JJA (Fig. 14~~6~~).

The Amazon basin experienced a record drought in 2010, which led to decreased vegetation greenness and a net carbon loss to the atmosphere (Xu et al., 2011; Gatti et al., 2014). Dry conditions in the north and center of the basin in the first three months were caused by the El Niño of late 2009-early 2010, and an enhanced and prolonged dry season in the southern areas of the basin was connected to an Atlantic sea surface temperature anomaly during the second half of the year (Gatti et al., 2014). According to our prior estimate, fire emissions minus NEP represented a near-zero net flux of -0.1 ± 2.1 Pg C y^{-1} in Jul-Sep 2010 (a period that includes peak drought conditions and fire counts of that year) and a sink of -1.9 ± 2.1 Pg C y^{-1} in Jul-Sep 2009 in the TC3 Tropical America region. (The fire emissions amounted to 2.0 Pg C y^{-1} and 0.2 Pg C y^{-1} in Jul-Sep 2010 and 2009, respectively, while NEP was 2.1 Pg C y^{-1} in both periods.) However, our GOSAT inversion suggests the reverse, -0.9 ± 0.6 Pg C y^{-1} vs. -0.4 ± 0.3 Pg C y^{-1} for Jul-Sep 2010 and 2009, respectively. (We do not report the analogous results for the in situ inversion, since the uncertainties are large in this undersampled region.) The prior estimate seems more consistent with the expected impact of drought on fluxes than the inversion estimate does. The inversion is hampered in the region by the relatively small number of GOSAT soundings that are retrieved and pass the quality filters, especially during the burning season (with substantial light scattering by aerosols) and the rainy season (with extensive cloud cover). The dearth of observations results in relatively large posterior uncertainties and/or sizable flux error correlations. Furthermore, there is differing data coverage, with 2010 having fewer observations than 2009 in the TC3 Tropical America region during the height of the fire season

926 (85 and 20 in Aug and Sep 2010 vs. 101 and 33 in 2009) and more observations than 2009 in
927 July (150 vs. 85). The differing data coverage itself could affect the flux estimates differently in
928 2009 and 2010. The Amazonica data set does not enable an evaluation of the flux estimates for
929 both 2009 and 2010, since the data set begins in 2010. However, comparison of the prior and
930 GOSAT model mole fractions in 2010 with the Amazonica data shows that biases for both can
931 vary substantially over time, e.g. in July vs. Aug-Sep (Fig. S1). This raises the possibility that
932 neither the prior nor the GOSAT inversion correctly estimates the interannual flux difference in
933 this region and also supports the idea that inversion bias can vary with data coverage.
934
935
936 **4. Discussion and conclusions**
937 We have successfully applied a ~~presented~~ global, high-resolution, batch Bayesian CO_2
938 inversion~~s~~ method ~~using~~ to surface in situ observations and passive satellite column
939 measurements from GOSAT ~~and in situ observations~~ and compared ~~them~~ the flux estimates with
940 ~~flux estimates~~ ones using Kalman filter and variational approaches that involve various
941 approximations. The exact inversion method provides full posterior error covariances, which
942 allows us to quantitatively evaluate the degree to which regional fluxes are constrained
943 independently of one another. However, for inversions over longer periods, using larger
944 volumes of data such as from OCO-2, or at higher flux resolution, more computationally
945 efficient methods are essential.
946 The GOSAT inversion is generally better constrained than the in situ inversion, with
947 smaller posterior regional flux uncertainties and correlations, except in places like North
948 America and northern and southern high-latitude ocean where the in situ observation networks

used provide relatively good coverage. Note that our in situ inversion did not make use of all the surface monitoring sites that operated during the analysis period, omitting for example a number of sites operated exclusively by agencies in Canada, Australia, and Europe (http://ds.data.jma.go.jp/gmd/wdcgg/cgi-bin/wdcgg/catalogue.cgi), and that the surface networks have been enhanced with additional sites since then. Furthermore, the in situ data sets that we used for evaluation of the inversions, including JR-STATION and Amazonica, could also be used as input in the inversions. And yet other aircraft data sets such as CONTRAIL, which samples large parts of the Pacific and some other areas (Niwa et al., 2012), and NOAA's regular aircraft profiles over mostly North America (https://www.esrl.noaa.gov/gmd/ccgg/aircraft/index.html) and column measurements such as from TCCON could be added. The use of GOSAT data in combination with in situ data provides even greater flux uncertainty reductions than the use of either data set alone, indicative of complementary constraints in the two datasets. Nevertheless, remaining coverage gaps ~~in GOSAT sampling~~, including a lack of GOSAT observations at high latitudes during winter over land and year-round over the ocean, and spatially, seasonally, and interannually varying coverage over tropical land, limit the ability to accurately resolve fluxes down to the scale of TransCom sub-continental/sub-ocean basin regions.

Our GOSAT inversion suggests a shift in the global terrestrial CO_2 sink from the tropics and south to the north, relative to the prior and the in situ inversion; for combined land and ocean fluxes, the GOSAT inversion produces a shift in the global sink from the tropics to the north and the south relative to the prior, and an increased source in the tropics of ~2 Pg C y^{-1} relative to the in situ inversion. Similar shifts are seen in studies using other inversion approaches, such as the inversion intercomparison of Houweling et al. (2015). This result may be driven at least in part

by sampling and uncorrected retrieval biases in the ACOS GOSAT data set, as suggested by sizable discrepancies between posterior mole fractions in the GOSAT-only inversion and surface in situ and lower-tropospheric HIPPO aircraft observations. While the shift in the global sink appears to be a robust feature of the inversions, the partitioning of the sink between land and ocean in the inversions using either in situ or GOSAT data is found to be sensitive to prior uncertainties because of negative correlations in the flux errors for the two domains. The loose prior uncertainties assumed in our baseline inversions may explain the larger ocean sink estimates compared to other studies, including CT2013B and the Houweling et al. (2015) intercomparison. A rationale for specifying loose prior uncertainties is that this allows the results to be driven more by the observations than by the prior estimates. However, in light of increasing confidence in estimates of the global ocean sink (e.g. from GCP), it may be more appropriate to start with a reliable set of ocean fluxes and apply tighter prior uncertainties similar to those from our sensitivity test. In any case, more weight should be given to combined land and ocean fluxes across latitudinal bands than to separate land and ocean flux estimates for the current observational configurations.

The GOSAT inversion indicates significantly less CO_2 uptake in summer of 2010 than in 2009 in the north, consistent with a previous GOSAT analysis and likely reflecting severe heat waves and drought especially across Eurasia. However, observations from the JR-STATION in situ network suggest that the GOSAT inversion (and to a lesser extent, the in situ + GOSAT inversion) exaggerates the 2010-2009 difference in uptake in Siberia, while the CASA-GFED prior reasonably estimates that quantity. Thus, it may not be accurate to assume that year-to-year posterior flux differences are insensitive to satellite retrieval biases, as was done in the other study. The prior, in situ posterior, and GOSAT posterior all indicate greater CO_2 uptake over

North America in spring to early summer of 2010 than in 2009, consistent with wetter conditions over large parts of the continent. Decreased net uptake in July-September of 2010 relative to 2009 in our prior appears to be consistent with record drought in the Amazon in 2010, while the GOSAT inversion shows the reverse. However, time-varying biases in both the prior model and the GOSAT inversion relative to Amazon aircraft profiles raise the possibility that neither one correctly estimates the interannual flux difference in this region and also support the idea that inversion bias can vary with data coverage. Overall, the results do demonstrate that climatic conditions can drive significant year-to-year variability in natural carbon fluxes on regional scales.

Gaps in coverage at higher latitudes, especially in winter, as well as limited sampling over tropical land are a fundamental limitation of passive satellite measurements (including OCO-2) and imply an important future role for active satellites such as NASA's proposed Active Sensing of CO_2 Emissions over Nights, Days, and Seasons (ASCENDS) mission (Kawa et al., 2010; ASCENDS Ad Hoc Science Definition Team, 2015). Ongoing development of thermal IR (TIR) CO_2 retrievals for GOSAT and the future GOSAT-2 with sensitivity to several layers from the lower troposphere to the lower stratosphere shows promise for producing sufficiently accurate data that could also help to fill NIR retrieval coverage gaps (Saitoh et al., 2017a; b). Additional in situ and TCCON measurements in ~~the~~ regions that are under-observed and challenging for forward model simulations~~tropics~~, especially ~~in~~ Africa, would also be valuable for ~~validating~~ improving bias corrections for satellite retrievals and evaluating flux inversions using satellite data.

~~The GOSAT inversion indicates significantly less CO_2 uptake in summer of 2010 than in 2009 in the north, consistent with a previous GOSAT analysis and likely reflecting severe heat~~

~~waves and drought especially across Eurasia. However, observations from the JR-STATION in situ network suggest that the GOSAT inversion (and to a lesser extent, the in situ + GOSAT inversion) exaggerates the 2010-2009 difference in uptake in Siberia, while the CASA-GFED prior reasonably estimates that quantity. Thus, it may not be accurate to assume that year-to-year posterior flux differences are insensitive to satellite retrieval biases, as was done in the other study. The prior, in situ posterior, and GOSAT posterior all indicate greater CO_2 uptake over North America in spring to early summer of 2010 than in 2009, consistent with wetter conditions over large parts of the continent. Decreased net uptake in July-September of 2010 relative to 2009 in our prior appears to be consistent with record drought in the Amazon in 2010, while the GOSAT inversion shows the reverse. However, time-varying biases in both the prior model and the GOSAT inversion relative to Amazon aircraft profiles raise the possibility that neither one correctly estimates the interannual flux difference in this region and also support the idea that inversion bias can vary with data coverage. Overall, the results do demonstrate that climatic conditions can drive significant year-to-year variability in natural carbon fluxes on regional scales.~~

~~This study has successfully applied the batch inversion method to satellite data at relatively high resolution to generate a solution useful for comparison with other techniques. However, for inversions over longer periods, using larger volumes of data such as from OCO-2, or at higher flux resolution, more computationally efficient methods are essential.~~

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work has been supported by the NASA Atmospheric CO_2 Observations from Space program element and the NASA Carbon Monitoring System Program. The NASA Goddard High-End Computing Program has provided access to and assistance with supercomputing resources at the NASA Center for Climate Simulation. The ACOS GOSAT data were produced by the ACOS/OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology using spectra acquired by the GOSAT Project. We thank Chris O'Dell for providing the ACOS data to us, John Miller and Manuel Gloor for their partnership in producing the Amazonica data (with support from NERC, FAPESP, ERC, NASA, CNPQ, NOAA, IPEN, and U. of Leeds for the Amazon Greenhouse measurement program led by L. Gatti), NOAA ESRL GMD CCGG for making their flask and continuous tower data publicly available, JMA (including Yukio Fukuyama and Atsushi Takizawa) for making their in situ data publicly available on the WDCGG website and providing assistance, and Steven Wofsy for making HIPPO data available. CT2013B results are provided by NOAA ESRL, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov. Many thanks go to Martha Butler for providing inversion code and documentation. We also thank Zhengxin Zhu for contributing to data processing, Liang Feng and Paul Palmer for their inversion region map, David Baker for advice on inversions, Sander Houweling for providing results from his intercomparison paper and for helpful discussion, Lesley Ott for help with using ACOS data files and for discussions, Chris O'Dell, Ed Dlugokencky, and especially Arlyn Andrews for comments on the manuscript, and Sourish Basu for discussions.

References

Andres, R. J., Boden, T. A., and Marland, G.: Monthly Fossil-Fuel CO_2 Emissions: Mass of Emissions Gridded by One Degree Latitude by One Degree Longitude, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A., doi:10.3334/CDIAC/ffe.MonthlyMass.2012, 2012.

Andrews, A. E., Kofler, J., Bakwin, P. S., Zhao, C., and Tans, P.: Carbon Dioxide and Carbon Monoxide Dry Air Mole Fractions from the NOAA ESRL Tall Tower Network, 1992-2009, Version: 2011-08-31, Path: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/in-situ/tower/~~ftp://ftp.cmdl.noaa.gov/ccg/towers/~~, 2009.

ASCENDS Ad Hoc Science Definition Team: Active Sensing of CO_2 Emissions over Nights, Days, and Seasons (ASCENDS) Mission Science Mission Definition Study (draft), https://cce.nasa.gov/ascends_2015/ASCENDS_FinalDraft_4_27_15.pdf, 2015.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P. and co-authors: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, Global Biogeochem. Cycles, 20, GB1002, doi:10.1029/2004GB002439, 2006.

Baldocchi, D., et al.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82, 2415–2434, 2001.

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, Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.

Belikov, D. A., Maksyutov, S., Krol, M., Fraser, A., Rigby, M., Bian, H. S., Agusti-Panareda, A., Bergmann, D., Bousquet, P., Cameron-Smith, P., Chipperfield, M. P., Fortems-Cheiney, A., Gloor, E., Haynes, K., Hess, P., Houweling, S., Kawa, S. R., Law, R. M., Loh, Z., Meng, L., Palmer, P. I., Patra, P. K., Prinn, R. G., Saito, R., and Wilson, C.: Off-line algorithm for calculation of vertical tracer transport in the troposphere due to deep convection, Atmos. Chem. Phys., 13, 1093-1114, doi:10.5194/acp-13-1093-2013, 2013.

Butler, M. P., Davis, K. J., Denning, A. S., and Kawa, S. R.: Using continental observations in global atmospheric inversions of CO_2: North American carbon sources and sinks, Tellus, 62B, 550–572, doi:10.1111/j.1600-0889.2010.00501.x, 2010.

Chatterjee, A. and Michalak, A. M.: Technical Note: Comparison of ensemble Kalman filter and variational approaches for CO_2 data assimilation, Atmos. Chem. Phys., 13, 11643–11660, doi:10.5194/acp-13-11643-2013, 2013.

Chevallier, F., et al.: What eddy-covariance measurements tell us about prior land flux errors in CO_2-flux inversion schemes, Global Biogeochem. Cycles, 26, GB1021, doi:10.1029/2010GB003974, 2012.

Chevallier, F., Palmer, P. I., Feng, L., Boesch, H., O'Dell, C. W., and Bousquet, P.: Toward robust and consistent regional CO_2 flux estimates from in situ and spaceborne measurements of atmospheric CO_2, Geophys. Res. Lett., 41, 1065–1070, doi:10.1002/2013GL058772, 2014~~3~~.

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO_2 fluxes: methods and perspectives, Climatic Change, 103:69–92, DOI 10.1007/s10584-010-9909-3, 2010.

Connor, B. J., Bösch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, J. Geophys. Res., 113, A05305, doi:10.1029/2006JD008336, 2008.

Crisp, D.: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2), Proc. SPIE 9607, Earth Observing Systems XX, 960702, doi:10.1117/12.2187291, 2015.

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.

Deng, F., Jones, D. B. A., O'Dell, C. W., Nassar, R., and Parazoo, N. C.: Combining GOSAT XCO_2 observations over land and ocean to improve regional CO_2 flux estimates, J. Geophys. Res. Atmos., 121, 1896–1913, doi:10.1002/2015JD024157, 2016.

Dlugokencky, E. J., Lang, P. M., Masarie, K. A., Crotwell, A. M., and Crotwell, M. J.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2012, Version: 2013-08-28, available at: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/ (last access: 18 February 2014), 2013.

Eldering, A., Wennberg, P. O., Crisp, D., Schimel, D. S., Gunson, M. R., Chatterjee, A., Liu, J., Schwandner, F. M., Sun, Y., O'Dell, C. W., Frankenberg, C., Taylor, T., Fisher, B., Osterman, G. B., Wunch, D., Hakkarainen, J., Tamminen, J., and Weir, B.: The Orbiting

Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes, Science, 358, eaam5745, doi:10.1126/science.aam5745, 2017.

Engelen, R. J., Denning, A. S., and Gurney, K. R.: On error estimation in atmospheric CO_2 inversions, J. Geophys. Res., 107, 4635, doi:10.1029/2002JD002195, 2002.

Enting, I. G. and Mansbridge, J. V.: Seasonal sources and sinks of atmospheric CO_2: Direct inversion of filtered data, Tellus, 41B, 111-126, 1989.

Enting, I. G., Trudinger, C. M., and Francey, R. J.: A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO_2, Tellus, 47B, 35–52, 1995.

Feng, L., Palmer, P. I., Boesch, H., and Dance, S.: Estimating surface CO_2 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., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R.: Estimates of European uptake of CO_2 inferred from GOSAT X_{CO2} retrievals: sensitivity to measurement bias inside and outside Europe, Atmos. Chem. Phys., 16, 1289–1302, doi:10.5194/acp-16-1289-2016, 2016.

Frankenberg, C., Kulawik, S. S., Wofsy, S., Chevallier, F., Daube, B., Kort, E. A., O'Dell, C., Olsen, E. T., and Osterman, G.: Using airborne HIAPER Pole-to-Pole Observations (HIPPO) to evaluate model and remote sensing estimates of atmospheric carbon dioxide, Atmos. Chem. Phys., 16, doi:10.5194/acp-2015-961, 2016.

Gatti, L. V., Gloor, M., Miller, J. B., Doughty, C. E., Malhi, Y., Domingues, L. G., Basso, L. S., Martinewski, A., Correia, C. S. C., Borges, V. F., Freitas, S., Braz, R., Anderson, L. O., Rocha, H., Grace, J., Phillips, O., and Lloyd, J.: Drought sensitivity of Amazonian carbon

balance revealed by atmospheric measurements, Nature, 506, 76-80, doi:10.1038/nature12957, 2014.

Gatti , L., Gloor, E., and Miller, J.: Greenhouse gas profile measurements (CO, CO_2, CH_4) above the forest canopy at four sites for the Amazonica project. NCAS British Atmospheric Data Centre, accessed 5 Dec 2016. http://catalogue.ceda.ac.uk/uuid/7201536a8b7a1a96de584e9b746acee3, 2016.

Giglio, L., Csiszar, I., and Justice, C. O.: Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, J. Geophys. Res., 111, G02016, doi:10.1029/2005JG000142, 2006.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, Biogeosciences, 7, 1171–1186, doi:10.5194/bg-7-1171-2010, 2010.

GLOBALVIEW-CO_2: Cooperative Atmospheric Data Integration Project-Carbon Dioxide, CD-ROM, NOAA-ESRL, Boulder, Colo. [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov,Path:ccg/co2/ GLOBALVIEW], 2009.

Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., Petron, G., Nehrkorn, T., Eluszkiewicz, J., Henderson, J., Wen, D., Lin, J., Fischer, M., Sweeney, C., and Michalak, A. M.: North American CO_2 exchange: intercomparison of modeled estimates with results from a fine-scale atmospheric inversion, Biogeosciences, 9, 457–475, 2012.

Guerlet, S., Basu, S., Butz, A., Krol, M., Hahne, P., Houweling, S., Hasekamp, O. P., and Aben, I.: Reduced carbon uptake during the 2010 Northern Hemisphere summer from GOSAT, Geophys. Res. Lett., 40, 2378–2383, doi:10.1002/grl.50402, 2013.

Gurney, K., Law, R., Rayner, P., and Denning, A. S.: TransCom 3 Experimental Protocol, Department of Atmospheric Science, Colorado State University, USA, Paper 707 (Available at http://transcom.colostate.edu/TransCom_3/transcom_3.html), 2000.

Gurney, K. R., et al.: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002.

~~Gurney, K. R., Chen, Y. H., Maki, T., Kawa, S. R., Andrews, A., and Zhu, Z.: Sensitivity of atmospheric CO_2 inversions to seasonal and interannual variations in fossil fuel emissions, J. Geophys. Res., 110, D10308, doi:10.1029/2004JD005373, 2005.~~

Hayes, D. J., Turner, D. P., Stinson, G., McGuire, A. D., Wei, Y., West, T. O., Heath, L. S., deJong, B., McConkey, B. G., Birdsey, R. A., Kurz, W. A., Jacobson, A. R., Huntzinger, D. N., Pan, Y., Post, W. M., and Cook, R. B.: Reconciling estimates of the contemporary North American carbon balance among terrestrial biosphere models, atmospheric inversions, and a new approach for estimating net ecosystem exchange from inventory-based data, Global Change Biology, doi: 10.1111/j.1365-2486.2011.02627.x, 2012.

Houborg, R., Rodell, M., Li, B., Reichle, R., and Zaitchik, B.: Drought indicators based on model assimilated GRACE terrestrial water storage observations, Wat. Resour. Res, 48, W07525, doi:10.1029/2011WR011291, 2012.

Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., Deng, F., Dlugokencky, E. J., Feng, L., Ganshin, A., Hasekamp, O., Jones, D., Maksyutov, S., Marshall, J., Oda, T., O'Dell, C. W., Oshchepkov, S., Palmer, P. I., Peylin, P., Poussi, Z., Reum, F., Takagi, H.,

Yoshida, Y., and Zhuravlev, R.: An intercomparison of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements, J. Geophys. Res. Atmos., 120, 5253–5266, doi:10.1002/2014JD022962, 2015.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703–4715, 2001.

Kawa, S. R., Erickson III, D. J., Pawson, S., and Zhu, Z.: Global CO_2 transport simulations using meteorological data from the NASA data assimilation system, J. Geophys. Res., 109, D18312, doi:10.1029/2004JD004554, 2004.

Kawa, S. R., Mao, J., Abshire, J. B., Collatz, G. J., Sun, X., and Weaver, C. J.: Simulation studies for a space-based CO_2 lidar mission, Tellus B, 62, 759-769, doi:10.1111/j.1600-0889.2010.00486.x, 2010.

Kim, J., Kim, H. M., Cho, C.-H., Boo, K.-O., Jacobson, A. R., Sasakawa, M., Machida, T., Arshinov, M., and Fedoseev, N.: Impact of Siberian observations on the optimization of surface CO_2 flux, Atmos. Chem. Phys., 17, doi:10.5194/acp-17-2881-2017, 2017.

Kulawik, S., Wunch, D., O'Dell, C., Frankenberg, C., Reuter, M., Oda, T., Chevallier, F., Sherlock, V., Buchwitz, M., Osterman, G., Miller, C. E., Wennberg, P. O., Griffith, D., Morino, I., Dubey, M. K., Deutscher, N. M., Notholt, J., Hase, F., Warneke, T., Sussmann, R., Robinson, J., Strong, K., Schneider, M., De Mazière, M., Shiomi, K., Feist, D. G., Iraci, L. T., and Wolf, J.: Consistent evaluation of ACOS-GOSAT, BESD-SCIAMACHY, CarbonTracker, and MACC through comparisons to TCCON, Atmos. Meas. Tech., 9, 683–709, doi:10.5194/amt-9-683-2016, 2016.

Law, R. M., et al.: TransCom model simulations of hourly atmospheric CO_2: Experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cycles, 22, GB3009, doi:10.1029/2007GB003050, 2008.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth Syst. Sci. Data, 5, 165–185, doi:10.5194/essd-5-165-2013, 2013.

Le Quéré, C., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F., Chini, L. P., Ciais, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Klein Goldewijk, K., Koven, C., Landschützer, P., Lauvset, S. K., Lefévre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., Pérez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Rödenbeck, C., Saito, S., Schuster, U., Schwinger, J., Séférian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S., and Zeng, N.: Global Carbon Budget 2015. Earth System Science Data, 7:349-396. doi:10.5194/essd-7-349-2015, 2015.

1241 Lindqvist, H., et al.: Does GOSAT capture the true seasonal cycle of carbon dioxide?, Atmos.
1242 Chem. Phys., 15, 13023–13040, doi:10.5194/acp-15-13023-2015, 2015.
1243 Liu, J., Bowman, K. W., Lee, M., Henze, D. K., Bousserez, N., Brix, H., Collatz, G. J.,
1244 Menemenlis, D., Ott, L., Pawson, S., Jones, D., and Nassar, R.: Carbon monitoring system
1245 flux estimation and attribution: Impact of ACOS-GOSAT XCO_2 sampling on the inference of
1246 terrestrial biospheric sources and sinks, Tellus B, 66, 22,486, doi:10.3402/tellusb.v66.22486,
1247 2014.
1248 Los, S. O., Collatz, G. J., Sellers, P. J., Malmström, C. M., Pollack, N. H., DeFries, R. S.,
1249 Bounoua, L., Parris, M. T., Tucker, C. J., and Dazlich, D. A.: A global 9-yr biophysical land
1250 surface dataset from NOAA AVHRR data. J. Hydrometeorol., 1, 183–199, 2000.
1251 Maksyutov, S., et al.: Regional CO_2 flux estimates for 2009–2010 based on GOSAT and ground-
1252 based CO_2 observations, Atmos. Chem. Phys., 13, 9351–9373, 2013.
1253 Nassar, R., Jones, D. B. A., Suntharalingam, P., Chen, J. M., Andres, R. J., Wecht, K. J.,
1254 Yantosca, R. M., Kulawik, S. S., Bowman, K. W., Worden, J. R., Machida, T., and
1255 Matsueda, H.: Modeling global atmospheric CO_2 with improved emission inventories and
1256 CO_2 production from the oxidation of other carbon species, Geosci. Model Dev., 3, 689–716,
1257 doi:10.5194/gmd-3-689-2010, 2010.
1258 Niwa, Y., Machida, T., Sawa, Y., Matsueda, H., Schuck, T. J., Brenninkmeijer, C. A. M., Imasu,
1259 R., and Satoh, M.: Imposing strong constraints on tropical terrestrial CO_2 fluxes using
1260 passenger aircraft based measurements, J. Geophys. Res., 117, D11303,
1261 doi:10.1029/2012JD017474, 2012.
1262 O'Dell, C. W., et al.: The ACOS CO_2 retrieval algorithm Part 1: Description and validation
1263 against synthetic observations, Atmos. Meas. Tech., 5, 99–121, 2012.

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

Orbe, C., Holzer, M., Polvani, L. M., and Waugh, D.: Air-mass origin as a diagnostic of tropospheric transport, J. Geophys. Res. Atmos., 118, 1459–1470, doi:10.1002/jgrd.50133, 2013.

Osterman, G., Eldering, A., Avis, C., O'Dell, C., Martinez, E., Crisp, D., Frankenberg, C., and Frankenberg, B.: ACOS Level 2 Standard Product Data User's Guide, v3.4, Jet Propulsion Laboratory, Pasadena, California, 2013.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S., and Hayes, D.: A Large and Persistent Carbon Sink in the World's Forests, Science, 333, 988-993, doi: 10.1126/science.1201609, 2011.

Parazoo, N. C., Denning, A. S., Kawa, S. R., Corbin, K. D., Lokupitiya, R. S., and Baker, I. T.: Mechanisms for synoptic variations of atmospheric CO_2 in North America, South America and Europe, Atmos. Chem. Phys., 8, 7239-7254, https://doi.org/10.5194/acp-8-7239-2008, 2008.

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S. R., Law, R. M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P. I., Prinn, R. G., Rigby, M., Saito, R., and Wilson, C.: TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in

the troposphere and lower stratosphere, Atmos. Chem. Phys., 11, 12813–12837, doi:10.5194/acp-11-12813-2011, 2011.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, PNAS, 104, 18925-18930, 2007.

Pinzon, J. E. and Tucker, C. J.: A non-stationary 1981-2012 AVHRR $NDVI_{3g}$ time series, Remote Sensing 6, 6929-6960; doi:10.3390/rs6086929, 2014.

Randerson, J. T., Thompson, M. V., and Malmstrom, C. M.: Substrate limitations for heterotrophs: Implications for models that estimate the seasonal cycle of atmospheric CO_2, Global Biogeochem. Cycles, 10(4), 585–602, doi:10.1029/96GB01981, 1996.

Rayner, P. J., Enting, I. G., Francey, R. J., and Langenfelds, R.: Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations, Tellus B, 51, 213–232, 1999.

Reuter, M., et al.: Satellite-inferred European carbon sink larger than expected, Atmos. Chem. Phys. Discuss., 14, 21,829–21,863, doi:10.5194/acpd-14-21829-2014, 2014.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Bovensmann, H., Burrows, J. P., Houweling, S., Liu, Y. Y., Nassar, R., Chevallier, F., Ciais, P., Marshall, J., and Reichstein, M.: How much CO_2 is taken up by the European terrestrial biosphere?, BAMS, 665-671, doi:10.1175/bams-d-15-00310.1, 2017.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S.,

Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, 2011.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific, Singapore, 2000.

Saeki, T., Maksyutov, S., Saito, M., Valsala, V., Oda, T., Andres, R. J., Belikov, D., Tans, P., Dlugokencky, E., Yoshida, Y., Morino, I., Uchino, O., and Yokota, T.: Inverse modeling of CO_2 fluxes using GOSAT data and multi-year ground-based observations, Sci. Online Lett. Atmos., 9, 45–50, doi:10.2151/sola.2013-011, 2013a.

Saeki, T., Maksyutov, S., Sasakawa, M., Machida, T., Arshinov, M., Tans, P., Conway, T. J., Saito, M., Valsala, V., Oda, T., Andres, R. J., and Belikov, D.: Carbon flux estimation for Siberia by inverse modeling constrained by aircraft and tower CO_2 measurements, J. Geophys. Res., 118, doi:10.1002/jgrd.50127, 2013b.

Saito, R., Patra, P. K., Sweeney, C., Machida, T., Krol, M., Houweling, S., Bousquet, P., Agusti-Panareda, A., Belikov, D., Bergmann, D., Bian, H. S., Cameron-Smith, P., Chipperfield, M. P., Fortems-Cheiney, A., Fraser, A., Gatti, L. V., Gloor, E., Hess, P., Kawa, S. R., Law, R. M., Locatelli, R., Loh, Z., Maksyutov, S., Meng, L., Miller, J. B., Palmer, P. I., Prinn, R. G., Rigby, M. and Wilson, C.: TransCom model simulations of methane: Comparison of vertical profiles with aircraft measurements, J. Geophys. Res., 118, 3891-3904, doi:10.1002/jgrd.50380, 2013.

Saitoh, N., Kimoto, S., Sugimura, R., Imasu, R., Shiomi, K., Kuze, A., Niwa, Y., Machida, T., Sawa, Y., and Matsueda, H.: Bias assessment of lower and middle tropospheric CO_2

concentrations of GOSAT/TANSO-FTS TIR version 1 product, Atmos. Meas. Tech., 10, 3877–3892, doi:10.5194/amt-10-3877-2017, 2017a.

Saitoh, N., Yamada, A., Itatsu, T., Imasu, R., Shiomi, K., and Niwa, Y.: Algorithm development for the TIR bands of GOSAT-2/TANSO-FTS-2: lessons from GOSAT/TANSO-FTS TIR CO_2 and CH_4 measurement, AGU Fall Meeting, New Orleans, US, 11-15 Dec. 2017, A33G-2472, 2017b.

Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., Davydov, D., Fofonov, A., Krasnov, O., Saeki, T., Koyama, Y., and Maksyutov, S.: Continuous measurements of methane from a tower network over Siberia, Tellus B: Chemical and Physical Meteorology, 62:5, 403-416, doi:-10.1111/j.1600-0889.2010.00494.x, 2010.

Sasakawa, M., Machida, T., Tsuda, N., Arshinov, M., Davydov, D., Fofonov, A., and Krasnov, O.: Aircraft and tower measurements of CO_2 concentration in the planetary boundary layer and the lower free troposphere over southern taiga in West Siberia: Long-term records from 2002 to 2011, J. Geophys. Res., doi:10.1002/jgrd.50755, 2013.

Schimel, D., Stephens, B. B., and Fisher, J. B.: Effect of increasing CO_2 on the terrestrial carbon cycle, PNAS, 112, doi/10.1073/pnas.1407302112, 2015.

Stephens, B. B., et al.: Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, Science, 316, 1732–1735, 2007.

Takagi, H., et al.: On the benefit of GOSAT observations to the estimation of regional CO_2 fluxes, Sci. Online Lett. Atmos., 7, 161–164, 2011.

Takagi, H., et al.: Influence of differences in current GOSAT XCO_2 retrievals on surface flux estimation, Geophys. Res. Lett., 41, 2598–2605, doi:10.1002/2013GL059174, 2014.

Takahashi, T., et al.: Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, Deep Sea Res., Part II, 56(8–10), 554–577, 2009.

Tarantola, A.: Inverse problem theory: methods for data fitting and model parameter estimation, Elsevier, Amsterdam, The Netherlands, 1987.

Tsutsumi, Y., Mori, K., Ikegami, M., Tashiro, T., and Tsuboi, K.: Long-term trends of greenhouse gases in regional and background events observed during 1998-2004 at Yonagunijima located to the east of the Asian continent, Atmospheric. Environment., 40, 5868-5879, 2006.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., and Arellano Jr., A. F.: Interannual variability in global biomass burning emissions from 1997 to 2004, Atmos. Chem. Phys., 6, 3423–3441, doi:10.5194/acp-6-3423-2006, 2006.

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.

van der Werf, G. R., Randerson, J. T., Giglio, L., van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., van Marle, M. J. E., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, Earth Syst. Sci. Data, 9, 697-720, doi:10.5194/essd-9-697-2017, 2017.

Wargan, K., Pawson, S., Olsen, M. A., Witte, J. C., Douglass, A. R., Ziemke, J. R., Strahan, S. E., and Nielsen, J. E.: The global structure of upper troposphere-lower stratosphere ozone in GEOS-5: A multiyear assimilation of EOS Aura data, J. Geophys. Res. Atmos., 120, 2013–2036, doi:10.1002/2014JD022493, 2015.

Wofsy, S. C., et al.: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philos. T. R. Soc. A, 369, 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

Wofsy, S. C., et al.: HIPPO Merged 10-second Meteorology, Atmospheric Chemistry, Aerosol Data (R_20121129), Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A. http://dx.doi.org/10.3334/CDIAC/hippo_010 (Release 20121129), 2012.

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmos. Chem. Phys., 11, 12317–12337, doi:10.5194/acp-11-12317-2011, 2011.

Xu, L., Samanta, A., Costa, M. H., Ganguly, S., Nemani, R. R., and Myneni, R. B.: Widespread decline in greenness of Amazonian vegetation due to the 2010 drought, Geophys. Res. Lett., 38, L07402, doi:10.1029/2011GL046824, 2011.

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., and Maksyutov, S.: Global concentrations of CO_2 and CH_4 retrieved from GOSAT: first preliminary results, SOLA, 5, 160–163, doi:10.2151/sola.2009-041, 2009.

1399 **Table 1.** Inversion Prior and Posterior Fluxes and Uncertainties Aggregated to TransCom 3 Regions, June 2009-May 2010.

TransCom Region	Prior		Fires	In Situ-Only			GOSAT-Only			In Situ + GOSAT			In Situ-Only, Tighter Prior			GOSAT-Only, Tighter Prior		
	Flux[a]	Unc	Flux	Flux	Unc	U.R. (%)[b]	Flux	Unc	U.R. (%)	Flux	Unc	U.R. (%)	Flux	Unc	U.R. (%)	Flux	Unc	U.R. (%)
Boreal North America	-0.1	0.6	0.1	0.1	0.1	81	0.2	0.3	43	0.1	0.1	87	-0.1	0.1	71	0.0	0.2	27
Temperate North America	-0.3	1.5	0.0	-0.6	0.1	93	-1.5	0.3	82	-0.7	0.1	96	-0.6	0.1	87	-1.2	0.2	71
Tropical America	0.4	1.0	0.1	-0.4	0.7	33	-0.2	0.2	79	-0.3	0.2	82	-0.2	0.3	26	-0.1	0.1	67
Temperate South America	0.4	1.2	0.1	0.4	0.8	31	1.1	0.2	85	1.0	0.2	85	0.3	0.3	27	0.9	0.1	73
Northern Africa	0.2	1.1	0.4	1.5	0.7	38	2.0	0.2	83	1.8	0.2	84	1.1	0.3	28	2.0	0.1	70
Southern Africa	0.0	1.2	0.8	-0.1	0.7	44	-0.6	0.1	89	-0.5	0.1	89	-0.1	0.3	38	-0.6	0.1	80
Boreal Asia	-0.1	1.2	0.1	-1.2	0.4	70	-0.5	0.4	65	-1.2	0.2	87	-1.0	0.2	60	-0.5	0.2	51
Temperate Asia	0.0	1.8	0.1	-0.1	0.7	61	1.4	0.4	79	0.9	0.3	85	-0.5	0.3	53	1.0	0.2	67
Tropical Asia	0.3	0.6	0.4	0.0	0.4	33	0.5	0.3	54	0.7	0.2	61	0.4	0.2	25	0.8	0.1	39
Australia	0.0	0.5	0.1	-0.2	0.4	15	0.6	0.2	71	0.3	0.1	73	-0.2	0.2	12	0.3	0.1	56
Europe	-0.1	1.3	0.0	0.6	0.4	70	-1.5	0.3	75	-0.6	0.2	87	0.3	0.2	61	-1.6	0.2	64
North Pacific Ocean	-0.5	0.3	0.0	-0.9	0.1	51	-0.5	0.2	29	-1.1	0.1	67	-0.8	0.1	29	-0.5	0.1	11
Tropical West Pacific Ocean	0.1	0.3	0.0	0.1	0.2	26	0.3	0.1	51	0.5	0.1	59	0.1	0.1	15	0.3	0.1	24
Tropical East Pacific Ocean	0.4	0.3	0.0	0.4	0.2	25	0.4	0.1	54	0.3	0.1	62	0.4	0.1	13	0.4	0.1	25
South Pacific Ocean	-0.3	0.6	0.0	-1.0	0.4	32	-1.1	0.3	51	-1.8	0.2	60	-0.6	0.2	18	-0.9	0.1	30
Arctic/Northern Ocean	-0.3	0.3	0.0	-0.4	0.1	56	-0.5	0.2	19	-0.1	0.1	62	-0.3	0.1	31	-0.4	0.1	5
North Atlantic Ocean	-0.2	0.2	0.0	-0.8	0.1	35	-0.5	0.1	23	-1.0	0.1	50	-0.5	0.1	12	-0.3	0.1	6
Tropical Atlantic Ocean	0.1	0.3	0.0	0.1	0.2	23	0.3	0.2	42	0.4	0.1	56	0.1	0.1	9	0.2	0.1	14
South Atlantic Ocean	-0.2	0.4	0.0	-0.5	0.3	19	-0.7	0.2	38	-1.0	0.2	49	-0.3	0.1	8	-0.5	0.1	18
Southern Ocean	-0.2	0.6	0.0	-0.4	0.3	48	-0.9	0.4	41	0.2	0.2	62	-0.5	0.1	34	-1.1	0.1	22
Tropical Indian Ocean	0.1	0.4	0.0	0.0	0.3	27	0.7	0.2	56	0.5	0.2	62	0.1	0.1	16	0.4	0.1	32
Southern Indian Ocean	-0.4	0.3	0.0	-0.5	0.2	15	-0.6	0.2	29	-0.6	0.2	40	-0.5	0.1	7	-0.4	0.1	11

[a]Fluxes in table, in Pg C, include fires but not fossil emissions

[b]Uncertainty reduction

Table 2. Normalized ~~Chi-Squared (~~Cost Function~~)~~ Values for the Inversions.

Inversion	A Priori	A Posteriori
In situ only	112.4	4.0
GOSAT only	2.2	0.8
In situ + GOSAT	12.2	1.1
In situ only, decreased prior uncertainties	112.4	5.0
GOSAT only, decreased prior uncertainties	2.2	0.8

Table 3. Mean 2010-2009 difference in mole fractions over June-July-August at Siberian sites (in ppm).

Site	Observations	Prior	GOSAT-Only Post	In Situ + GOSAT Post	Prior - Obs	(GOSAT-Only) - Obs	(In Situ + GOSAT) - Obs
VGN	5.2	5.3	7.4	6.6	0.1	2.2	1.4
AZV	7.0	6.3	8.1	7.1	-0.7	1.1	0.1
SVV	2.6	4.0	3.4	4.6	1.4	0.8	2.0
IGR	4.9	5.7	5.1	4.6	0.8	0.2	-0.3
KRS	6.6	5.4	3.8	3.2	-1.2	-2.8	-3.4
YAK	2.1	2.5	4.2	2.5	0.4	2.1	0.4

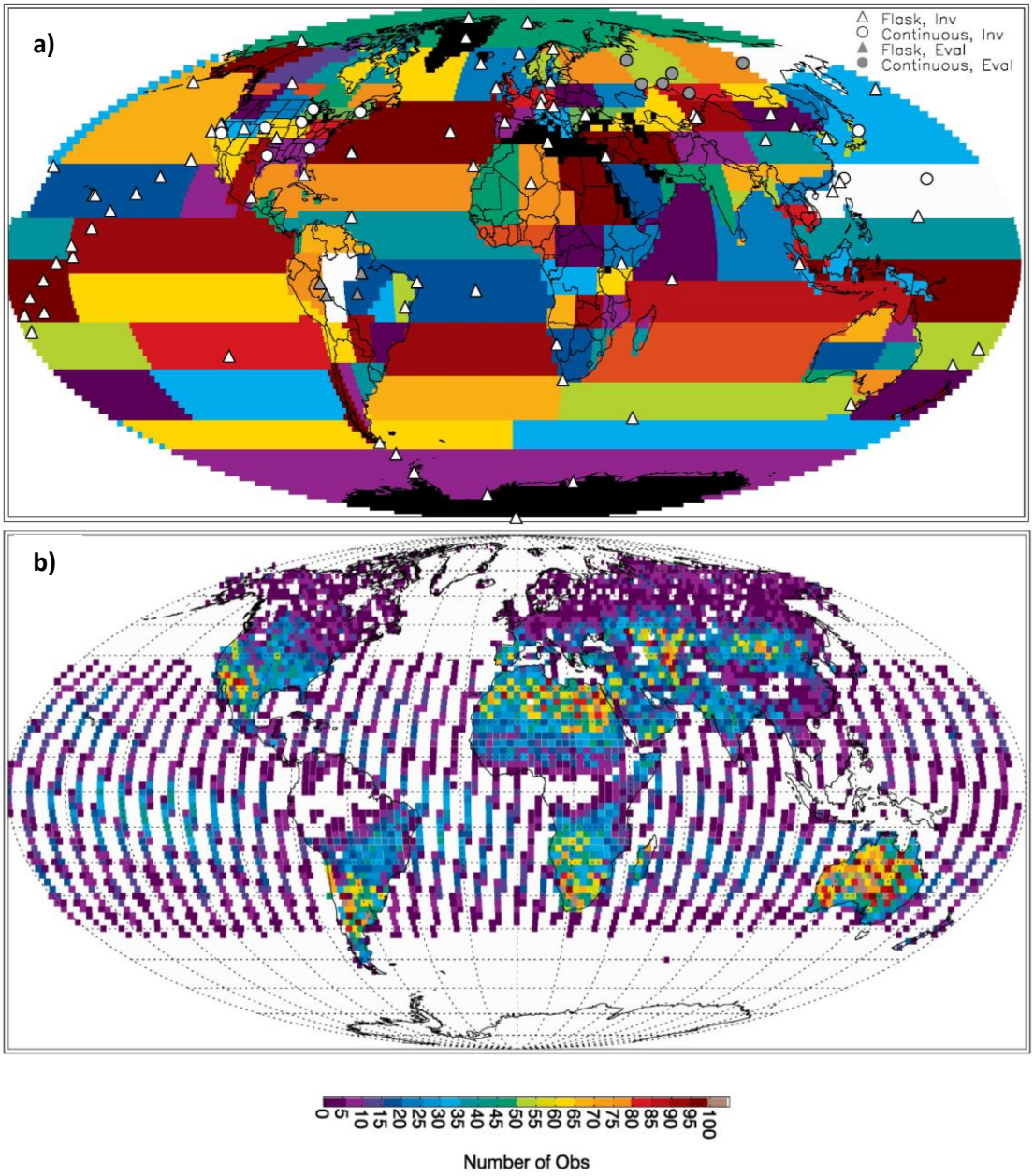


Figure 1. Locations of a) in situ observation sites and b) GOSAT XCO_2 observations used in the inversions. Also shown in a) are the 108 flux regions. Flask and continuous measurement sites ~~Triangles~~ in a) are represented by different symbols, ~~indicate flask sites, circles indicate continuous measurement sites~~and sites used in inversions and in their evaluation are represented by different colors. Observations in b) correspond to the ACOS B3.4 retrieval, are filtered and

1412 averaged over each hour and 2° x 2.5° PCTM model grid column, and are shown for June 2009-

1413 May 2010.

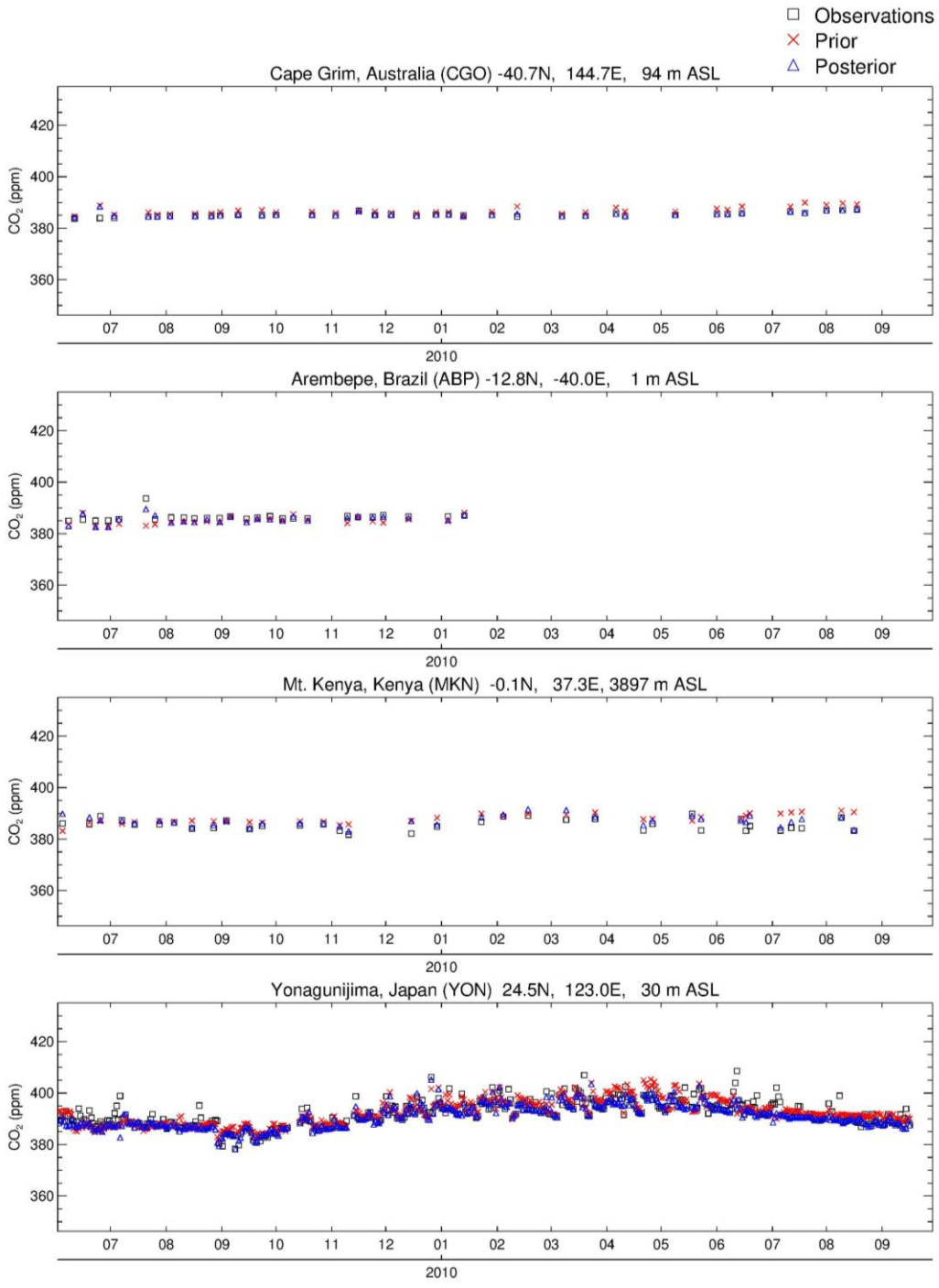
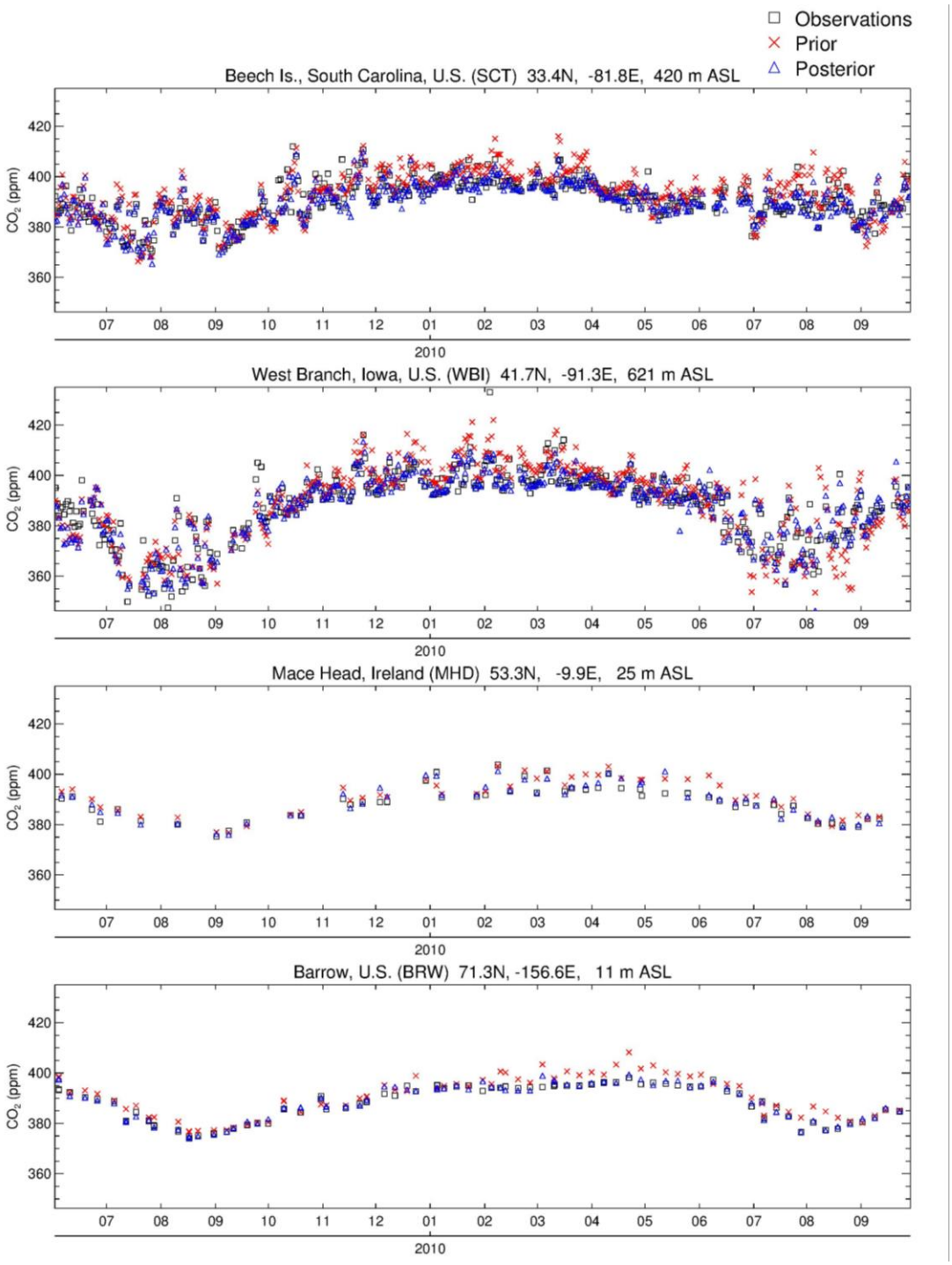


Figure 2. Comparison of model and observed time series of CO_2 mole fractions at selected surface sites. Posterior mole fractions are for the in situ-only inversion. Sites are arranged from south to north. Elevations include intake heights on towers where applicable.



1418

1419 **Figure 2.** (continued)

1420

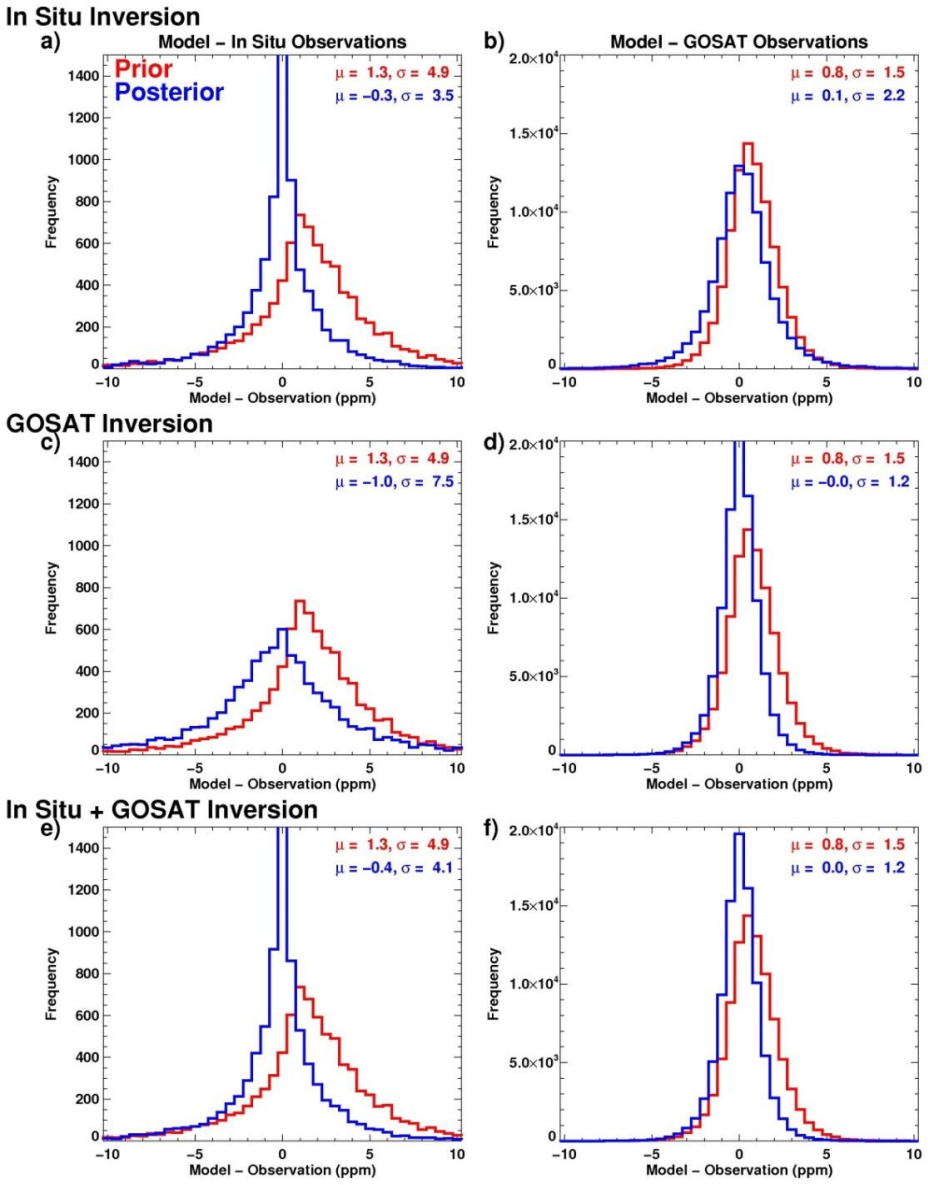


Figure 3. Full comparison of model and observations. Model-observation difference histograms are shown for (a) in situ-only inversion and in situ observations, (b) in situ-only inversion and GOSAT observations, (c) GOSAT-only inversion and in situ observations, (d) GOSAT-only inversion and GOSAT observations, (e) in situ + GOSAT inversion and in situ observations, and (f) in situ + GOSAT inversion and GOSAT observations. Mean differences and standard deviations are indicated in the panels.

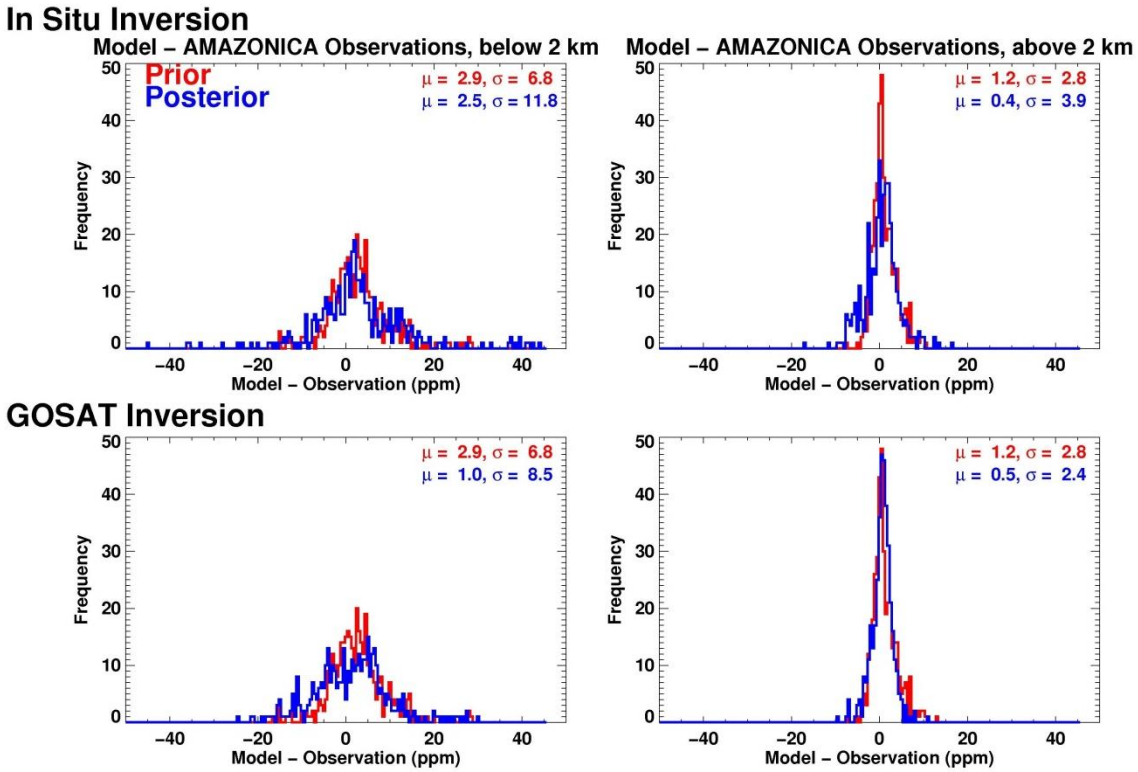


Figure 4. Comparison of model and Amazon aircraft observations (Amazonica project) over the period of overlap, Jan.-Sep. 2010. Top two panels show model-observation difference histograms for the in situ-only inversion and bottom two panels show results for the GOSAT-only inversion. Comparisons are shown separately for model and data below 2 km altitude (left) and above 2 km (right). Mean differences and standard deviations are indicated in the panels.

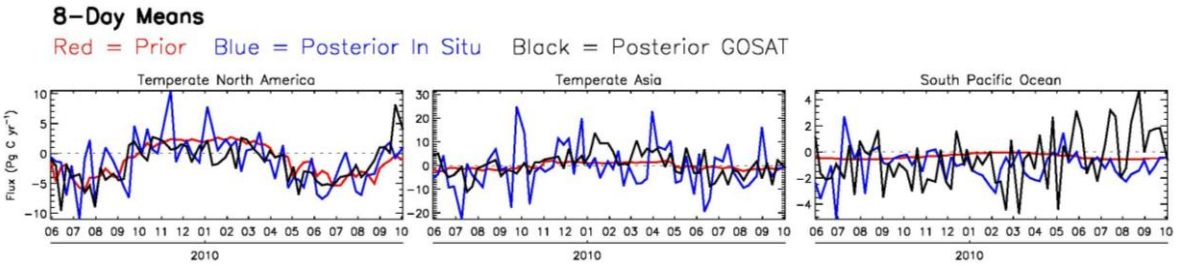
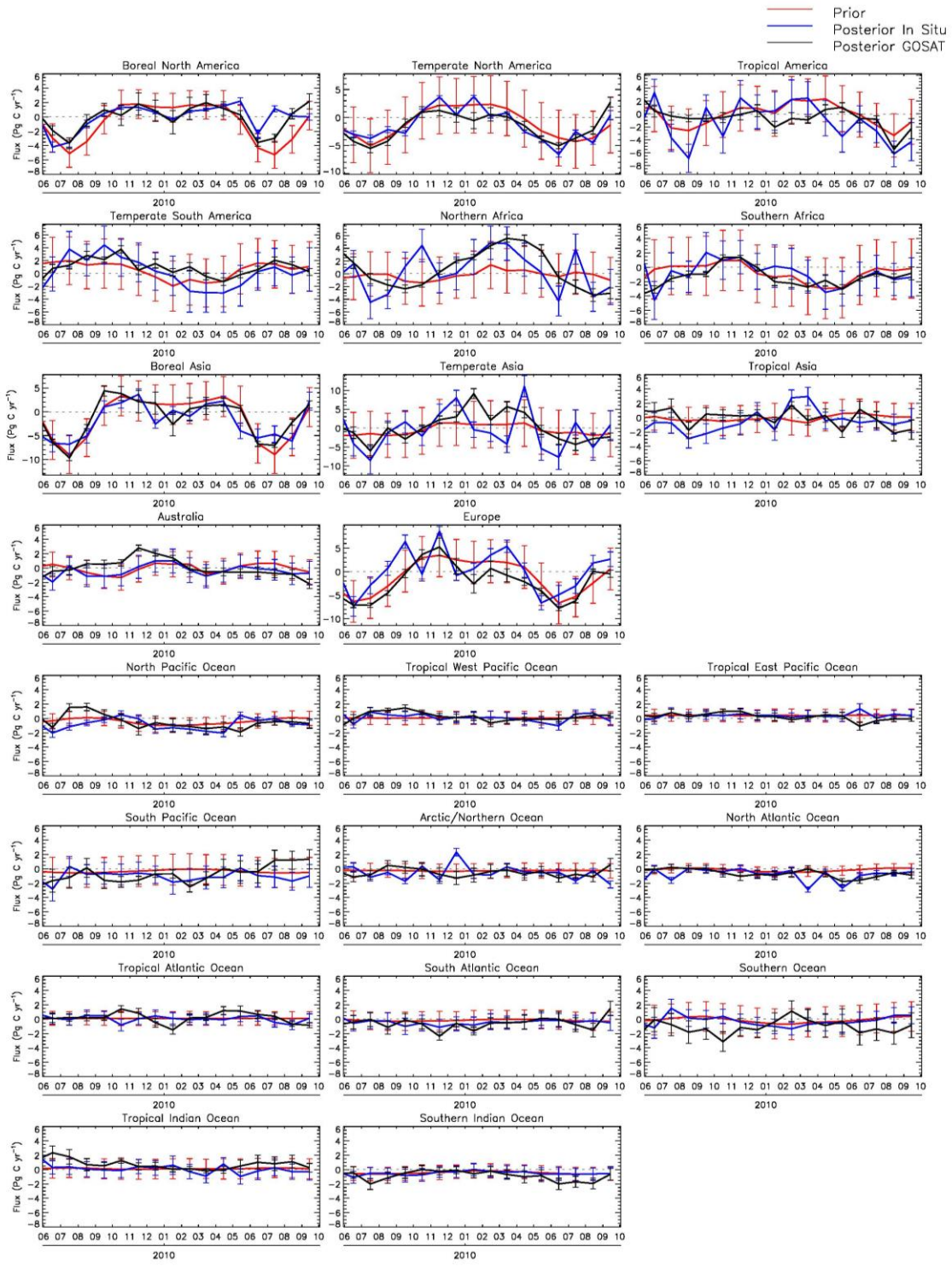


Figure 5. Prior, posterior in situ-only, and posterior GOSAT-only 8-day mean NEP (× -1) and ocean fluxes, aggregated over selected TransCom regions. Note that vertical scales are different in each of the panels.



1440

1441 **Figure 6.** Same as Fig. 5, except showing monthly means of fluxes for all TransCom regions,

1442 with error bars that represent 1σ uncertainties.

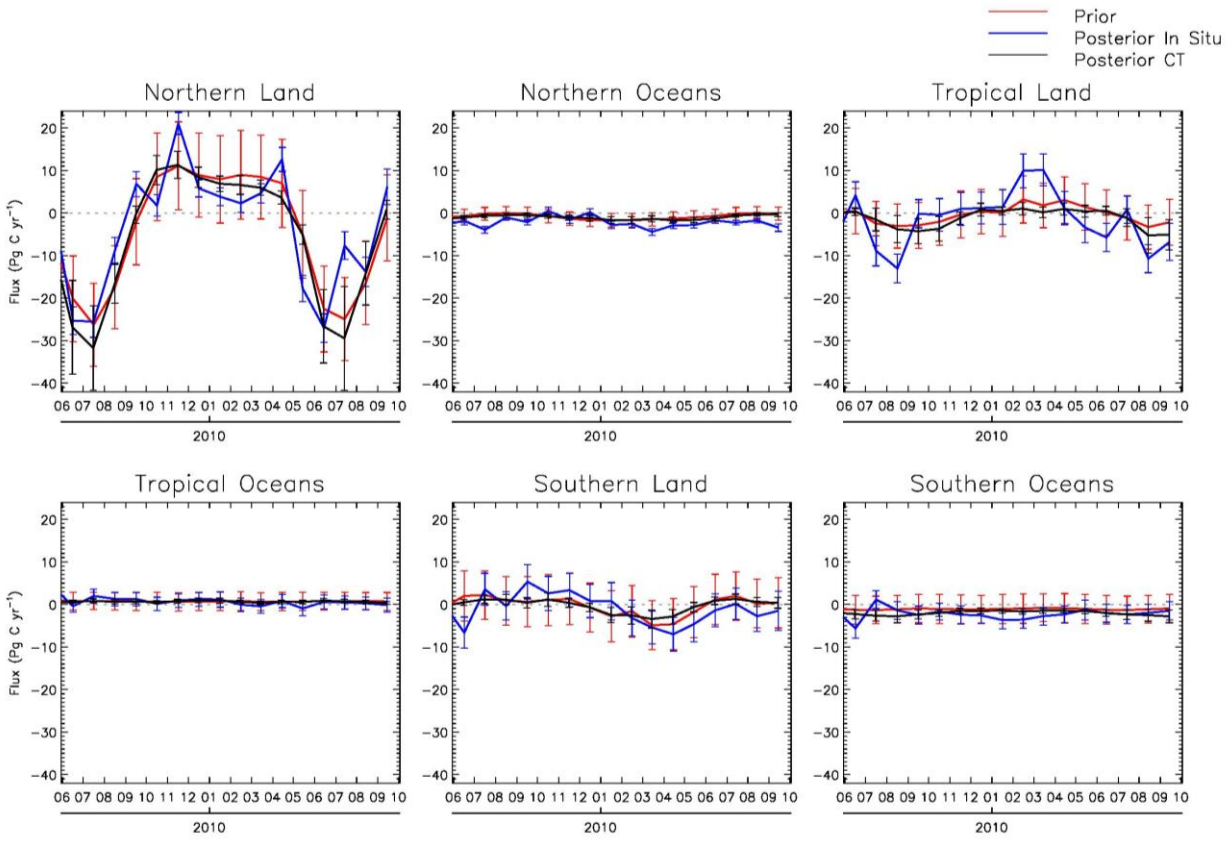
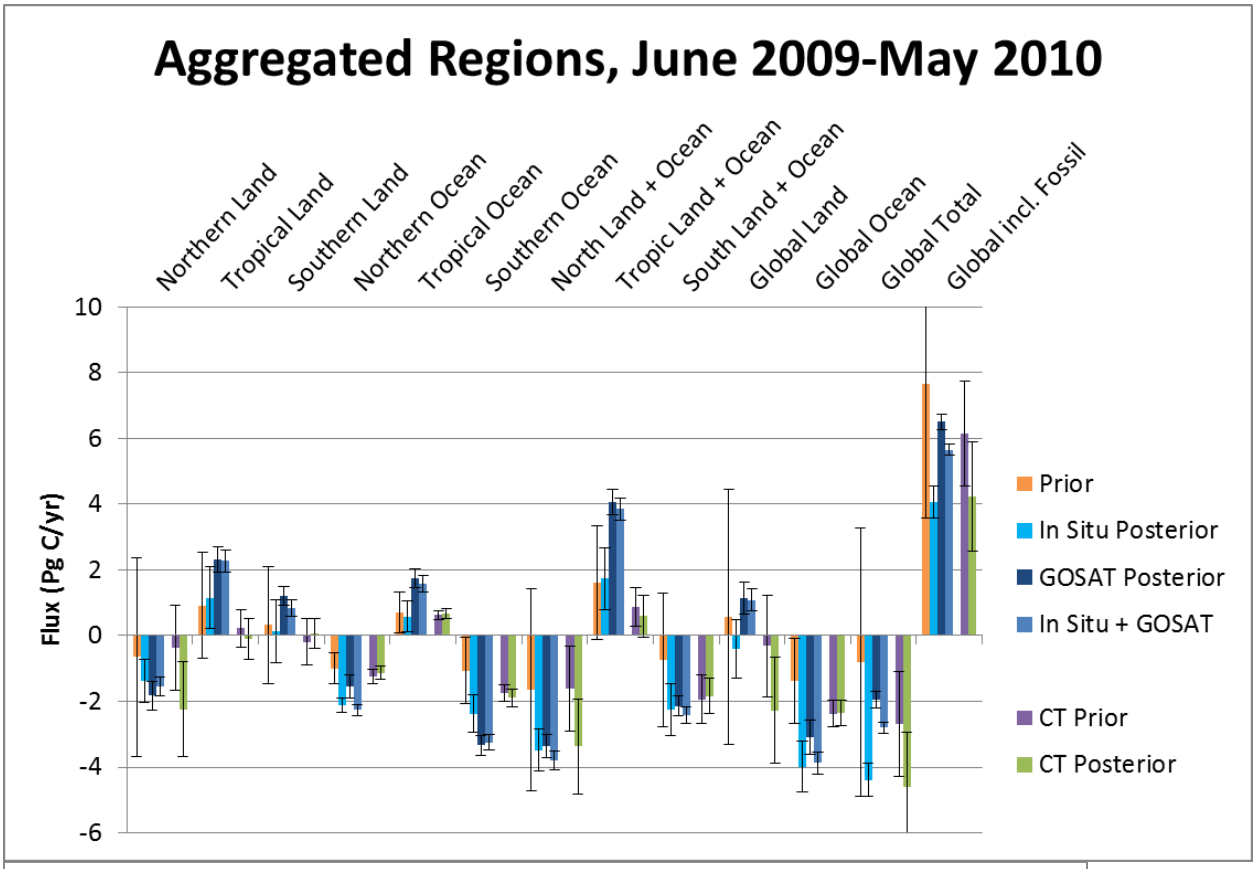
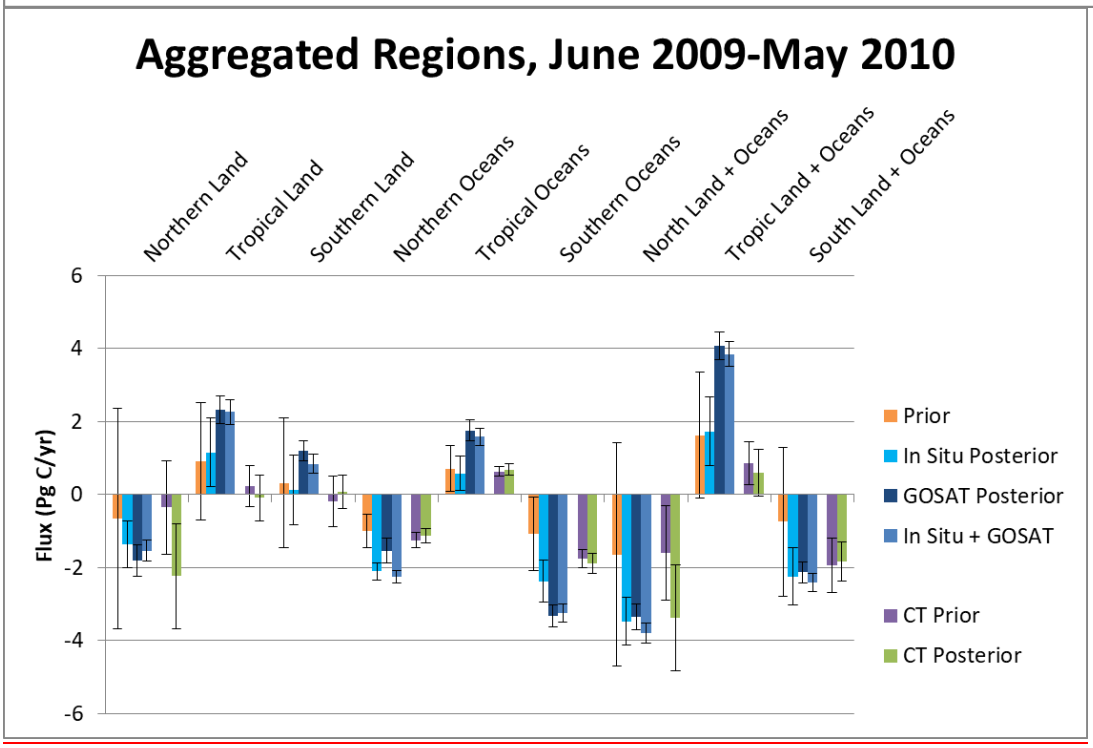


Figure 7. Comparison of our in situ-only inversion monthly mean NEP (× -1) and ocean fluxes, aggregated over large regions (as defined in TC3), with posterior fluxes from NOAA's CarbonTracker (CT2013B) data assimilation system. The priors shown are from our analysis; CT2013B priors are similar. Error bars represent 1σ uncertainties.


Aggregated Regions, June 2009-May 2010
Northern Land
Tropical Land
Southern Land
Northern Ocean
Tropical Ocean
Southern Ocean
North Land + Ocean
Tropic Land + Ocean
South Land + Ocean
Global Land
Global Ocean
Global Total
Global incl. Fossil
Flux (Pg C/yr)
10
8
6
4
2
0
-2
-4
-6
Prior
In Situ Posterior
GOSAT Posterior
In Situ + GOSAT
CT Prior
CT Posterior

1450


Aggregated Regions, June 2009-May 2010
Northern Land
Tropical Land
Southern Land
Northern Oceans
Tropical Oceans
Southern Oceans
North Land + Oceans
Tropic Land + Oceans
South Land + Oceans
Flux (Pg C/yr)
6
4
2
0
-2
-4
-6
Prior
In Situ Posterior
GOSAT Posterior
In Situ + GOSAT
CT Prior
CT Posterior

1451

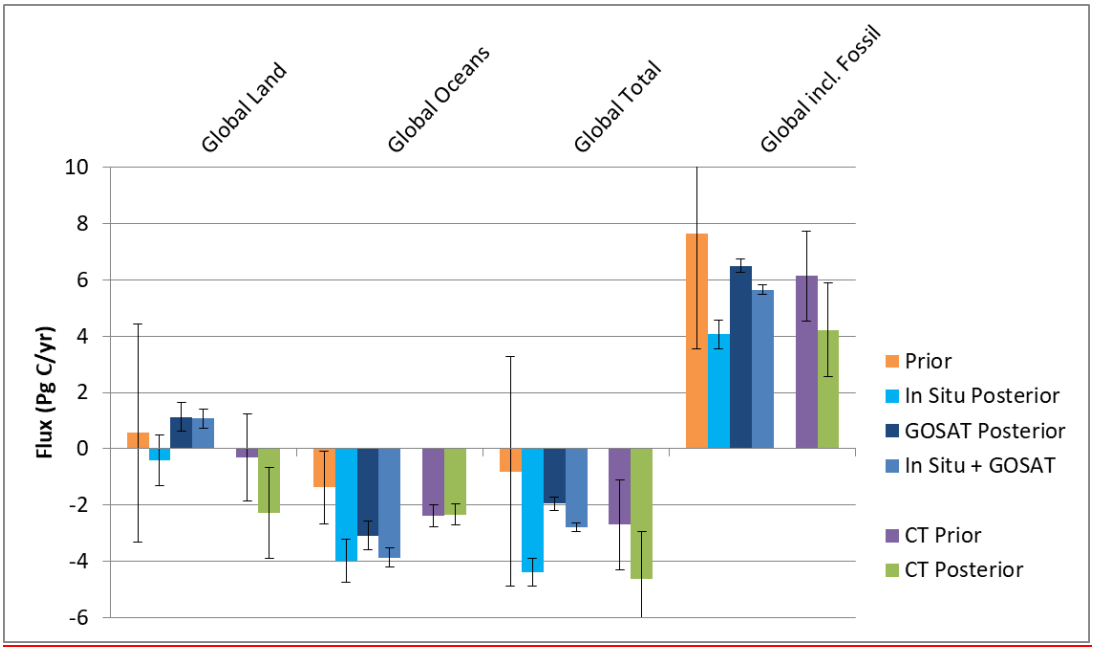


Figure 8. Twelve-month mean NEP (× -1), fire, and ocean fluxes aggregated over large regions. Included are results for the in situ-only, GOSAT-only, and in situ + GOSAT inversions as well as priors. Shown for comparison are priors and posteriors from CT2013B. Error bars represent 1σ uncertainties; for CT2013B, "external" (across a set of priors) as well as "internal" (within a particular inversion) uncertainties are included. In summing monthly CT2013B fluxes over the 12 months, we assumed zero error correlation between months.

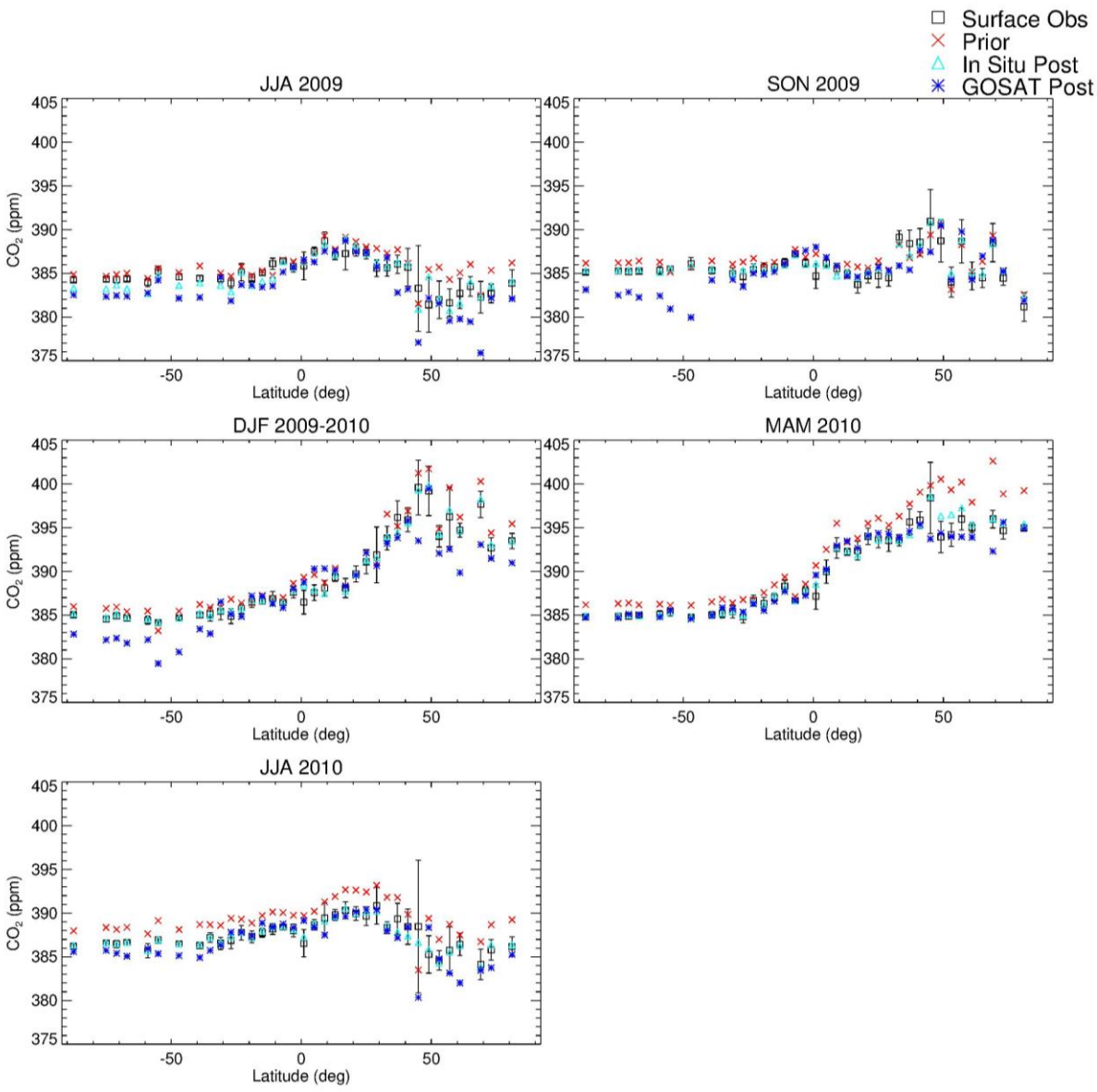
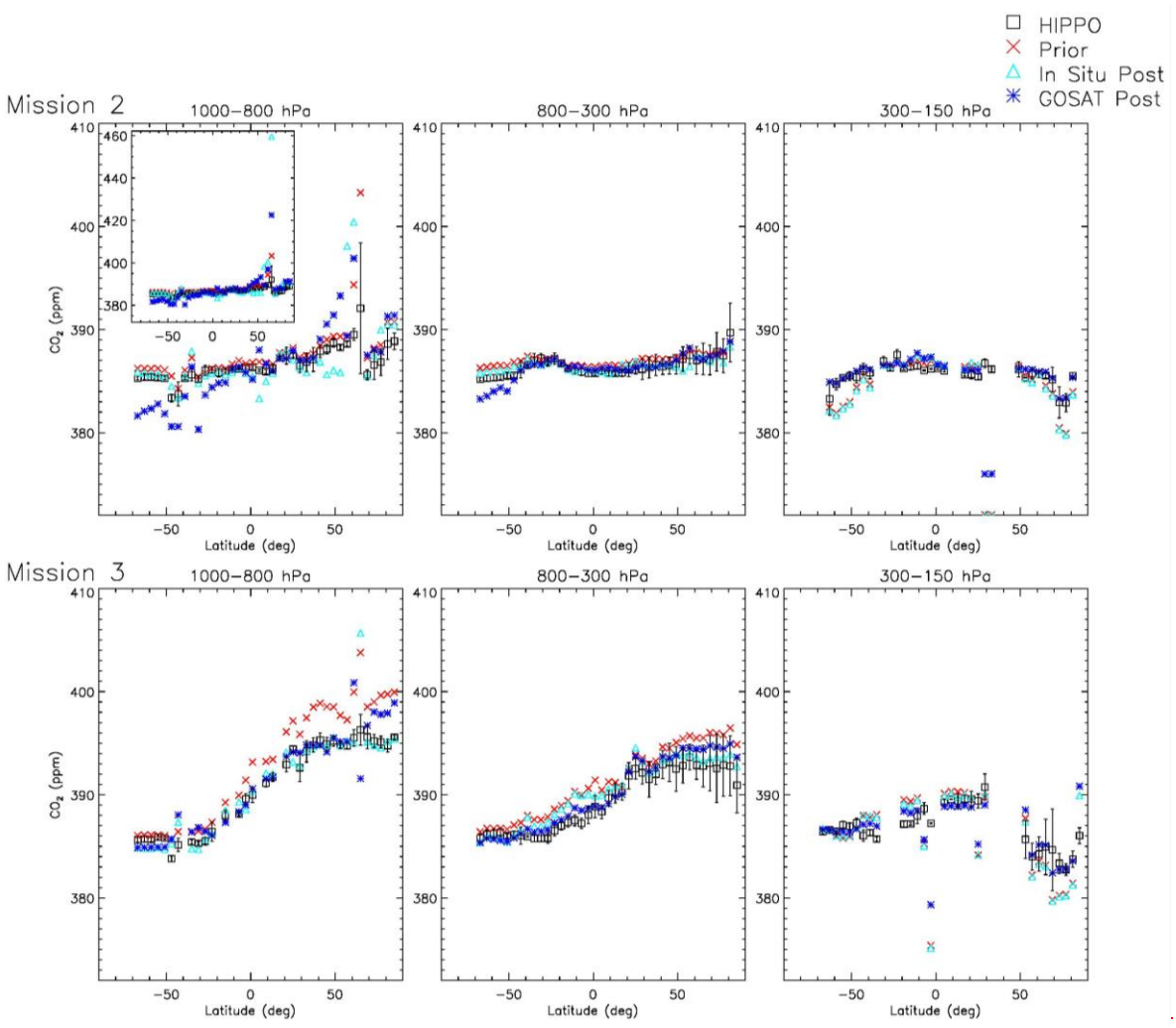


Figure 9. Latitudinal profiles of seasonal mean CO_2 mole fractions at surface sites for observations, prior, in situ-only posterior, and GOSAT-only posterior. Values are averaged in 4° bins. Error bars account for the spread of the observations within each season and bin as well as the uncertainty of each observation.


HIPPO
Prior
In Situ Post
GOSAT Post
Mission 2
Mission 3
1000–800 hPa
800–300 hPa
300–150 hPa
CO2 (ppm)
Latitude (deg)

1465

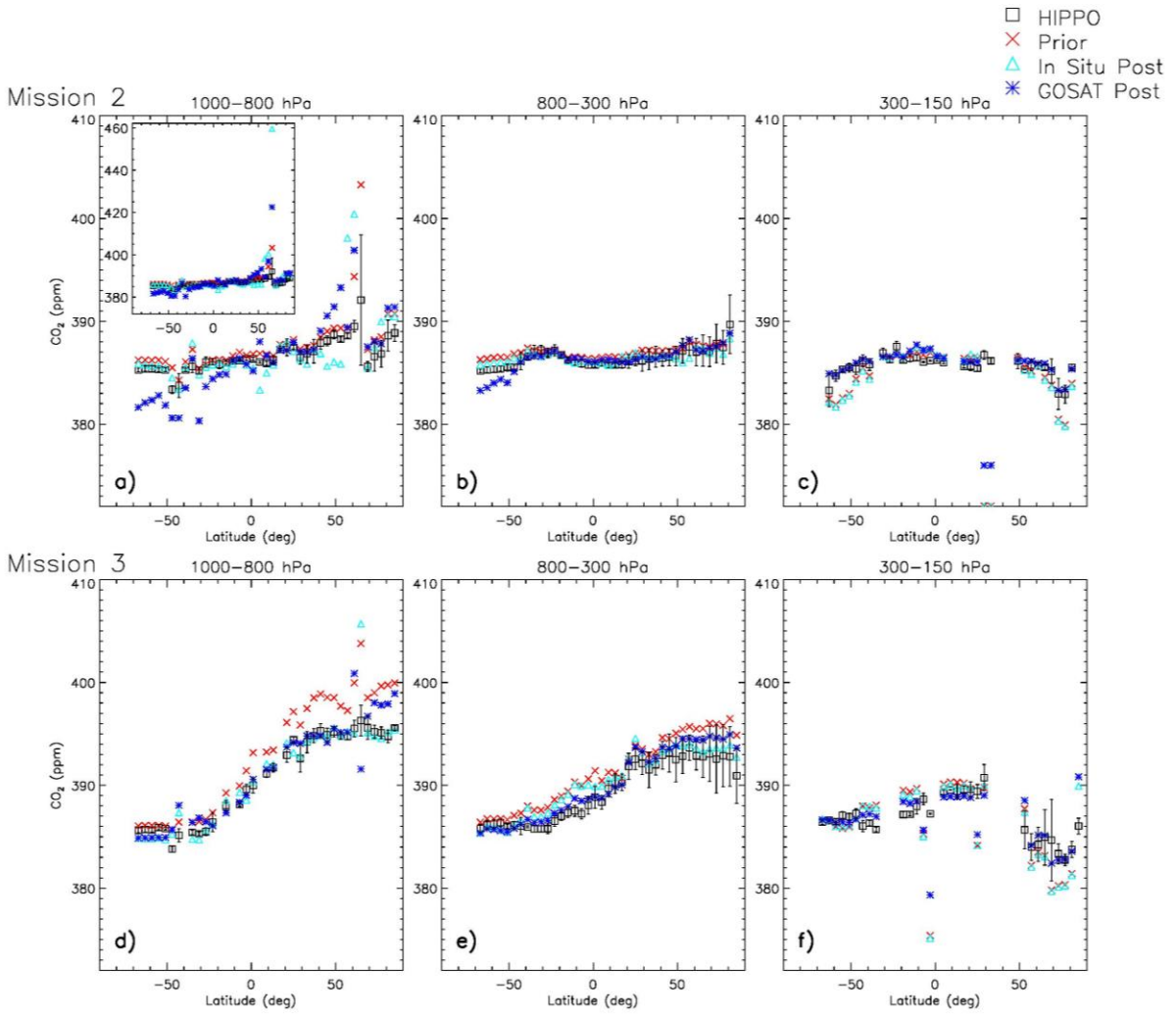
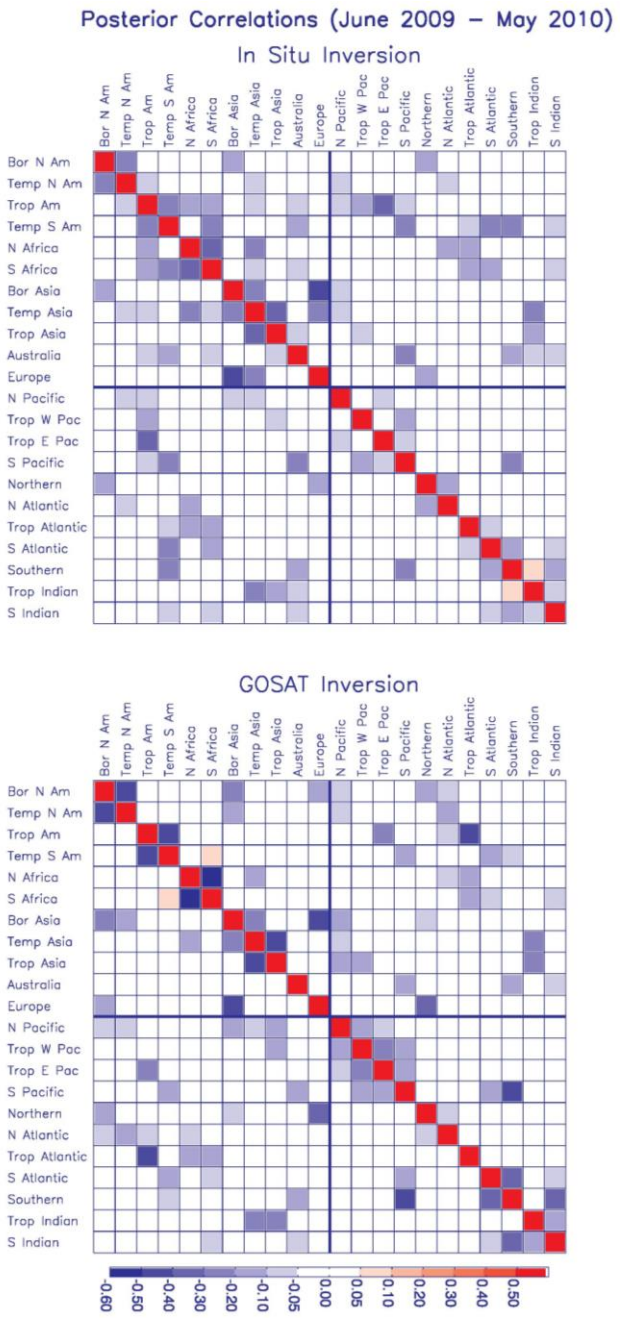


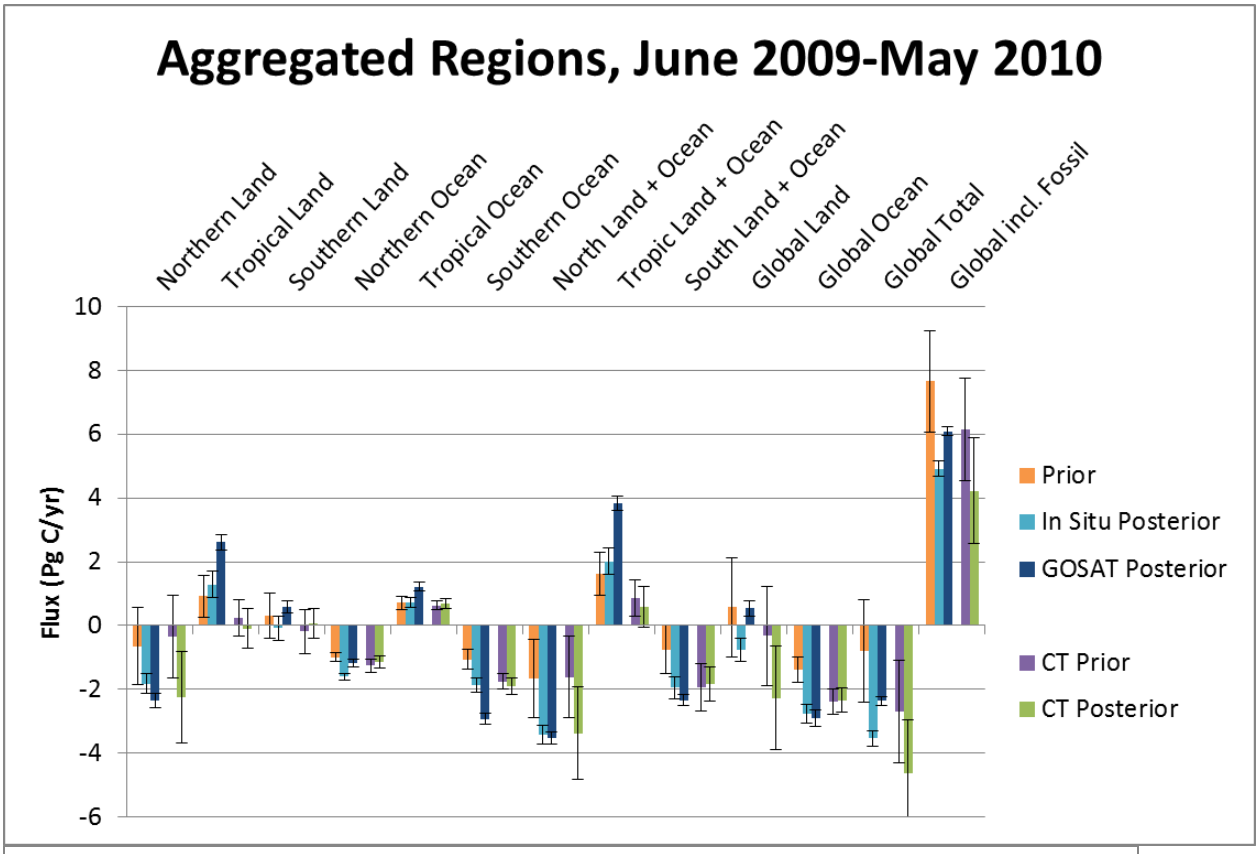
Figure 10. Latitudinal profiles of CO_2 mole fractions for HIPPO observations and co-sampled prior, in situ-only posterior, and GOSAT-only posterior. Mission 2 took place during Oct 31-Nov 22, 2009; Mission 3 took place Mar 24-Apr 16, 2010. Values are averaged in three altitude bins and 4° latitude bins. The inset in the first panel contains an expanded y-axis range that shows two points that do not fit into the default range. Flight segments over the temperate North American continent (east of -130°) are excluded from this comparison in order to focus on the Pacific. Error bars represent the standard deviations of the observations within each bin.



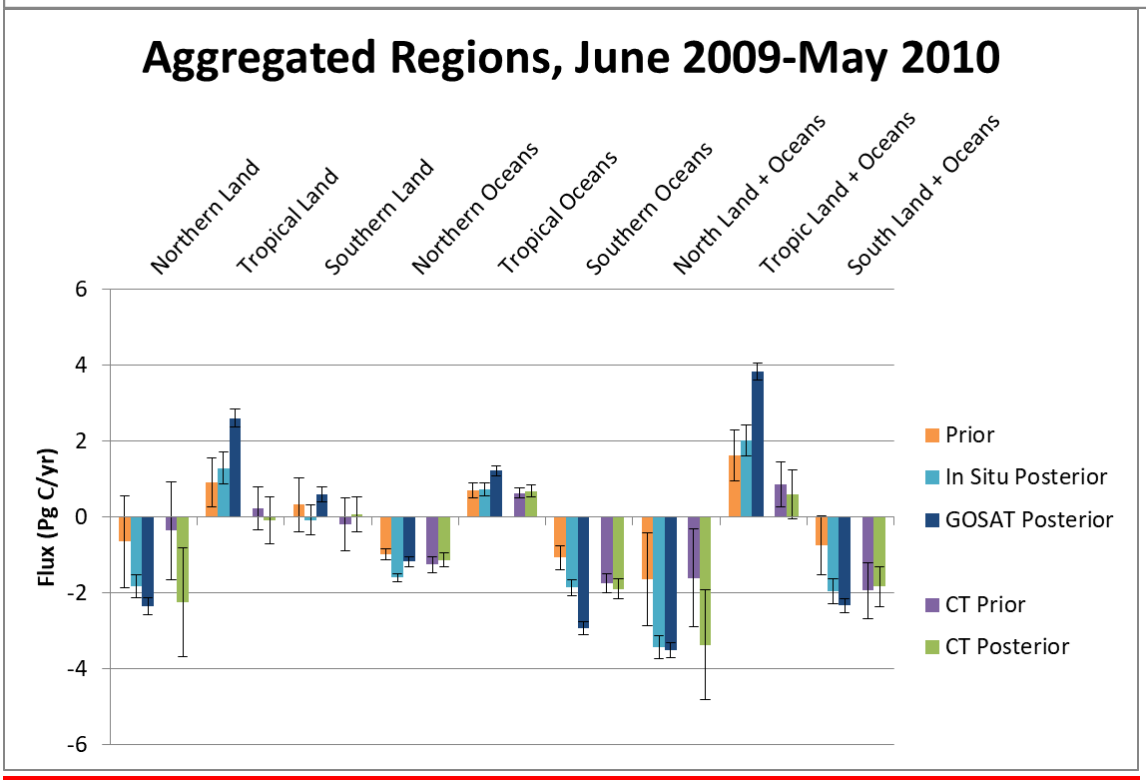
1474

1475 **Figure 11.** Posterior flux error correlations, aggregated to TC3 regions and a 12-month period,
1476 for (a) the in situ-only inversion, and (b) the GOSAT-only inversion. The correlation is equal to
1477 the error covariance divided by the product of the corresponding flux uncertainties (σ). Values
1478 on the main diagonal are equal to 1.

1479


Aggregated Regions, June 2009-May 2010
Northern Land
Tropical Land
Southern Land
Northern Ocean
Tropical Ocean
Southern Ocean
North Land + Ocean
Tropic Land + Ocean
South Land + Ocean
Global Land
Global Ocean
Global Total
Global incl. Fossil
Flux (Pg C/yr)
10
8
6
4
2
0
-2
-4
-6
Prior
In Situ Posterior
GOSAT Posterior
CT Prior
CT Posterior

1480


Aggregated Regions, June 2009-May 2010
Northern Land
Tropical Land
Southern Land
Northern Oceans
Tropical Oceans
Southern Oceans
North Land + Oceans
Tropic Land + Oceans
South Land + Oceans
Flux (Pg C/yr)
6
4
2
0
-2
-4
-6
Prior
In Situ Posterior
GOSAT Posterior
CT Prior
CT Posterior

1481

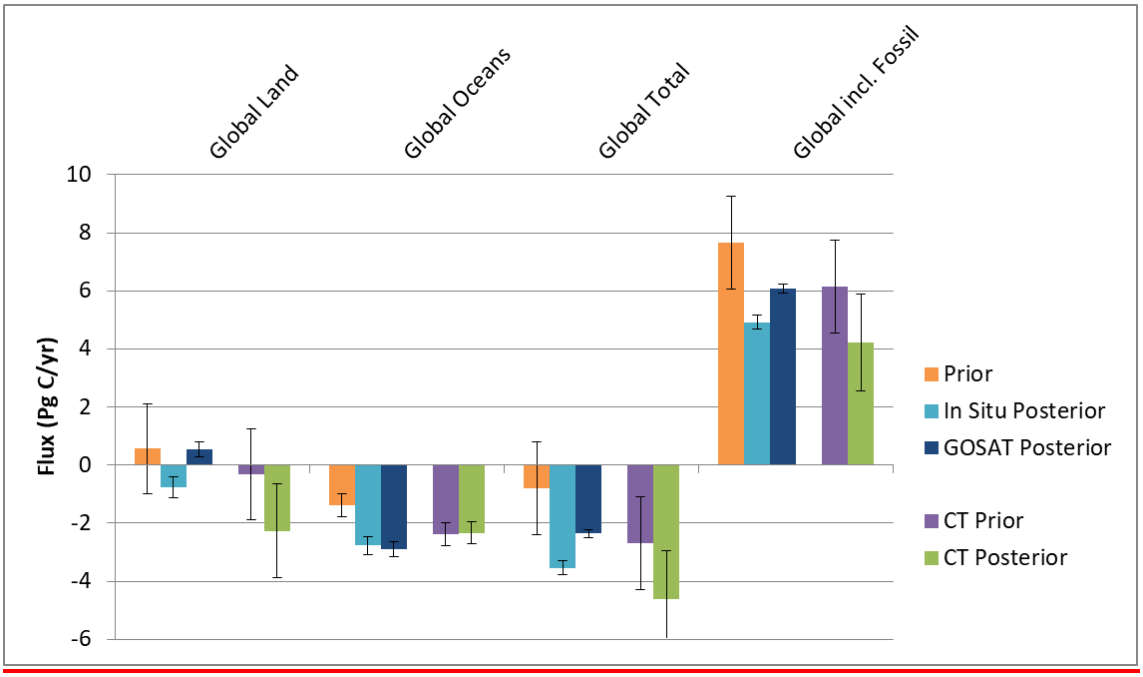
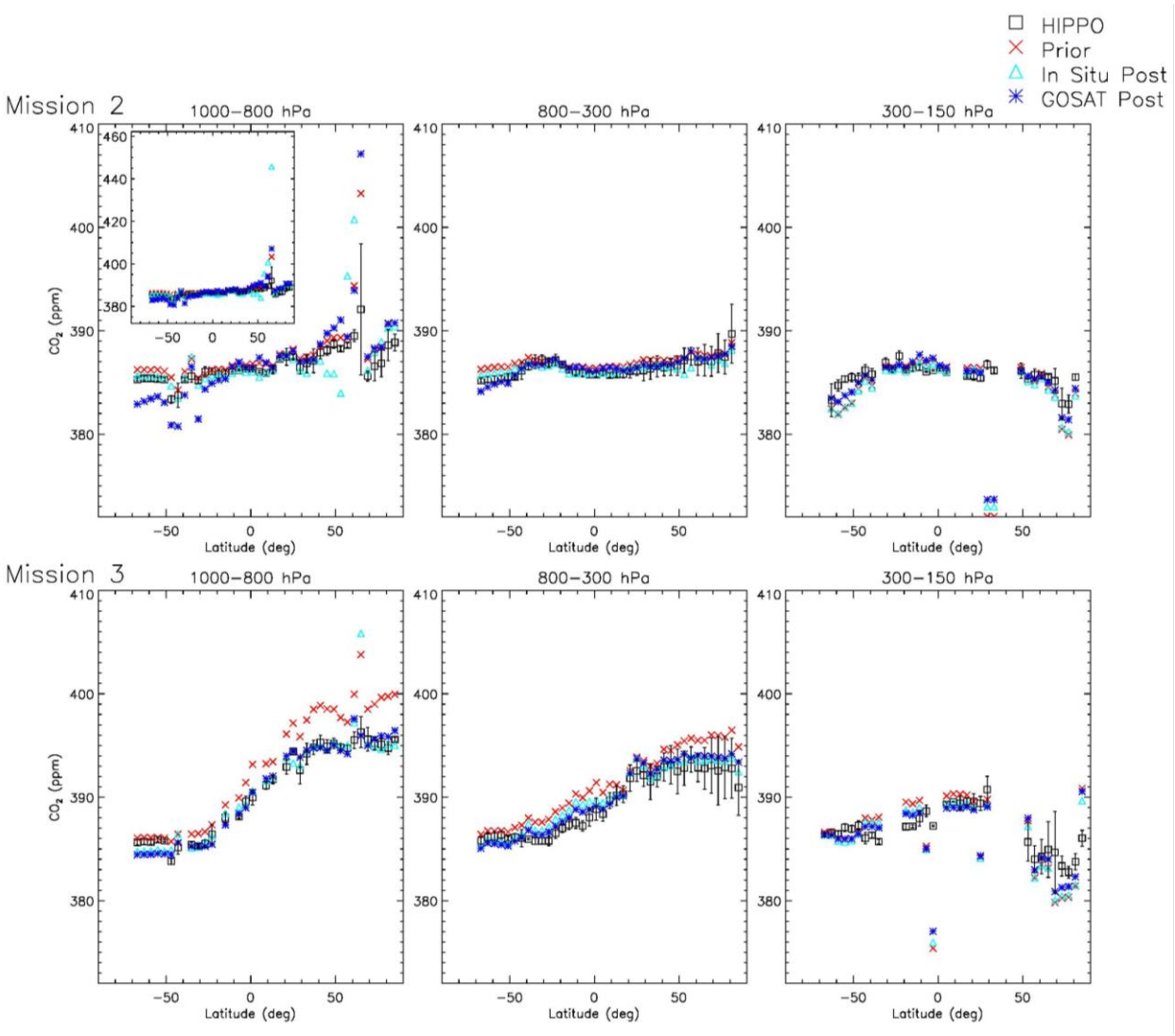
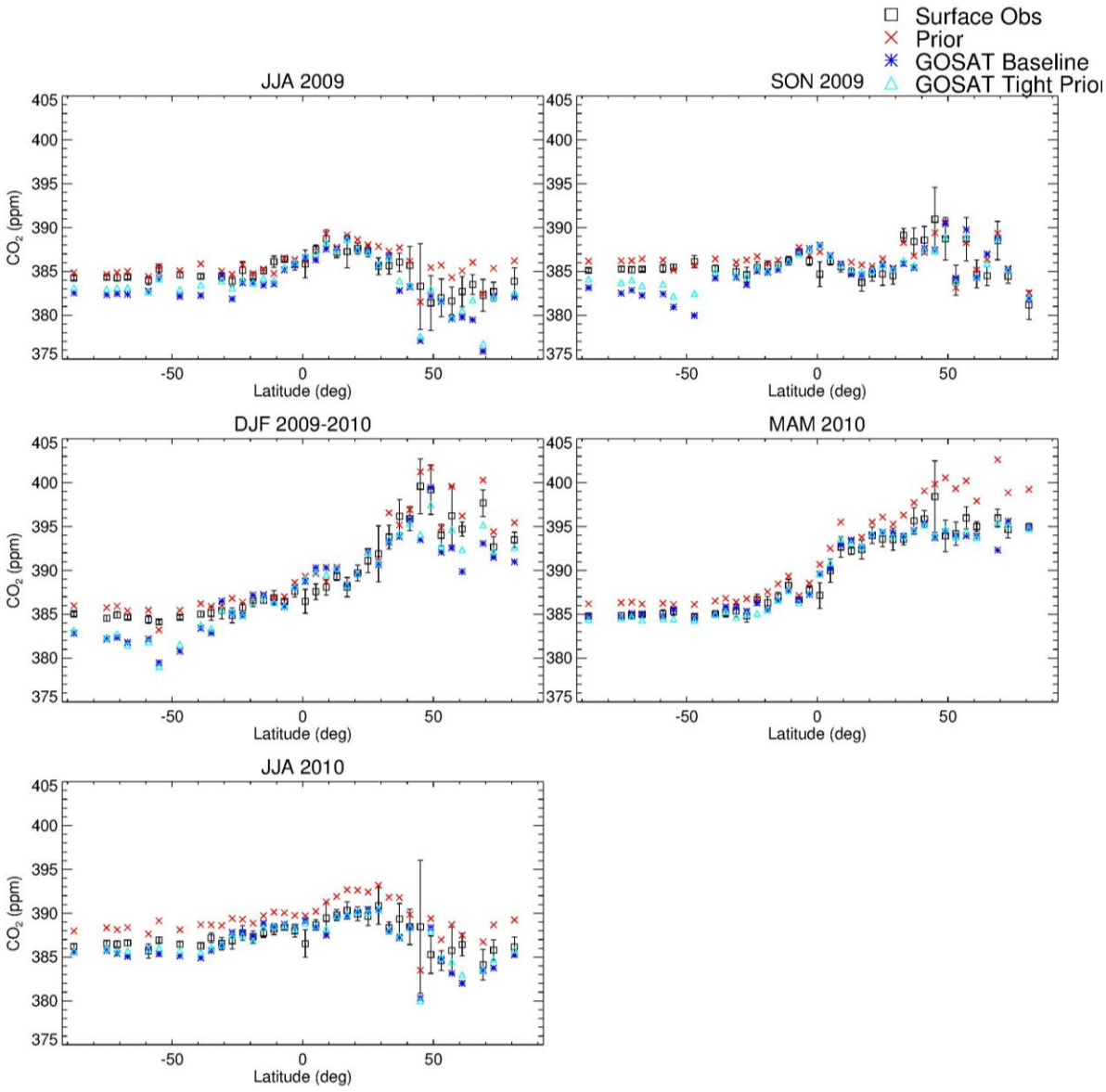


Figure 12. Similar to Fig. 8, except showing results for inversions with tighter prior constraints (with prior uncertainties similar to CarbonTracker's). Included are results for the in situ-only and GOSAT-only inversions. CT2013B results shown in Fig. 8 are repeated here. Error bars represent 1σ uncertainties.



~~Figure 13.~~ ~~Same as Fig. 10 except showing inversions with tighter prior uncertainties.~~



~~Figure 14.~~ ~~Same as Fig. 9 except showing GOSAT-only inversions with baseline vs. tighter prior uncertainties.~~

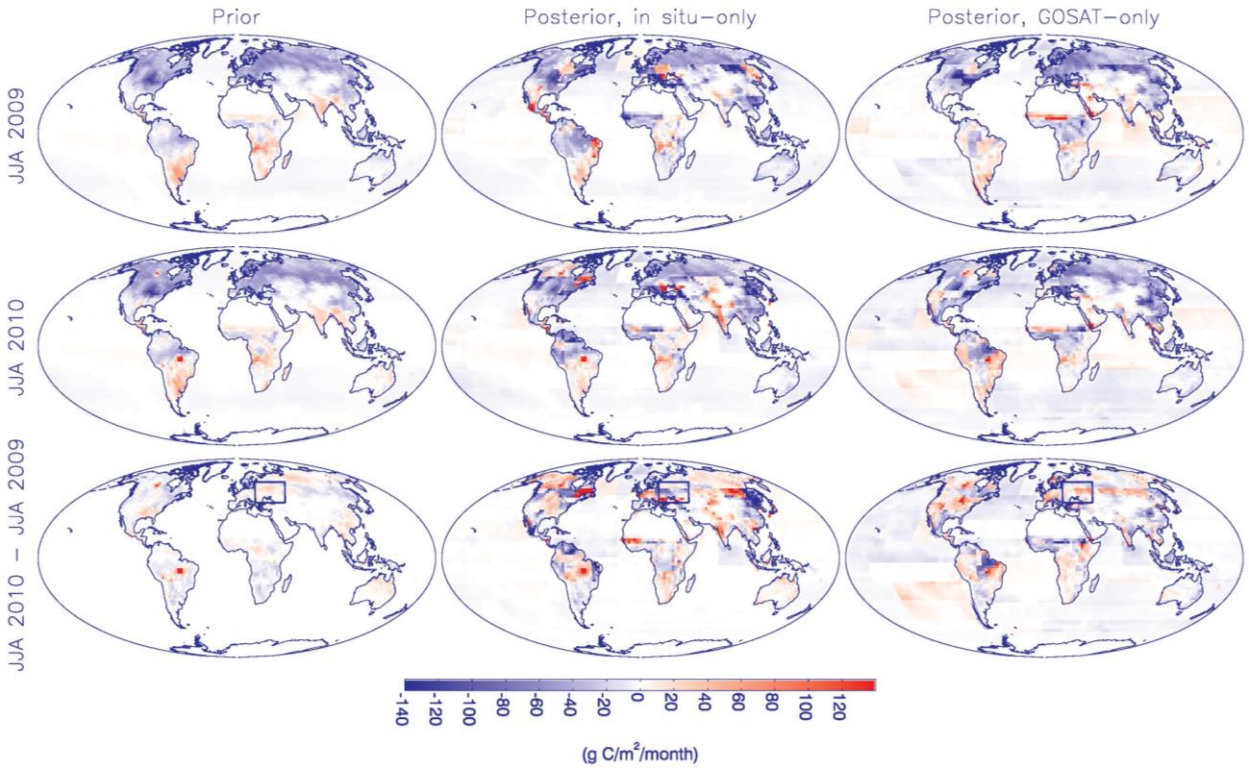


Figure 135. Comparison of spatial distribution of fluxes for June-July-August of 2010 vs. 2009. Included are natural and fire fluxes. Shown are fluxes for 2009 (top), 2010 (middle), and the 2010-2009 difference (bottom), for the priors (left), in situ-only inversion (middle), and GOSAT-only inversion (right). In the bottom row, boxes enclose the region around western Russia where there were intense heat waves, severe drought, and extensive fires. Note that the grid-scale spatial variability shown is not optimized in the inversions, so only patterns at the scale of the 108 flux regions contain information from the observations.

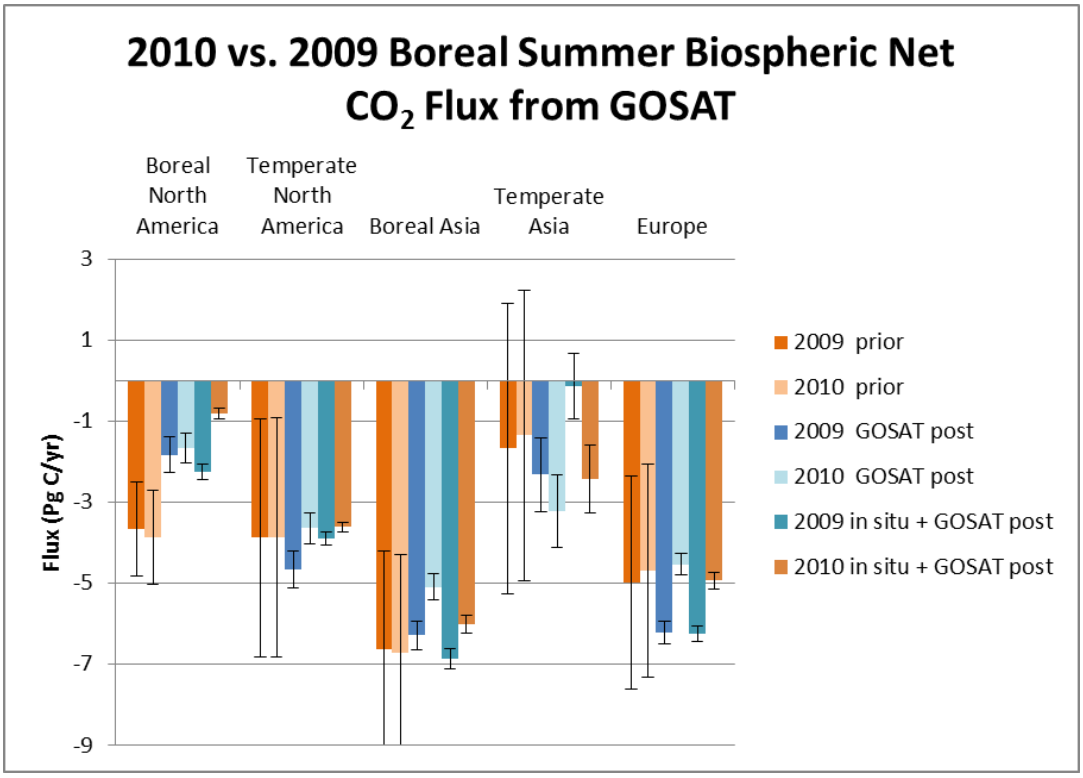


Figure 1~~4~~6. Comparison of prior, GOSAT-only posterior, and in situ + GOSAT posterior fluxes aggregated over northern regions for June-July-August of 2010 vs. 2009. Included are NEP (× -1) and fire fluxes. Error bars represent 1σ uncertainties.

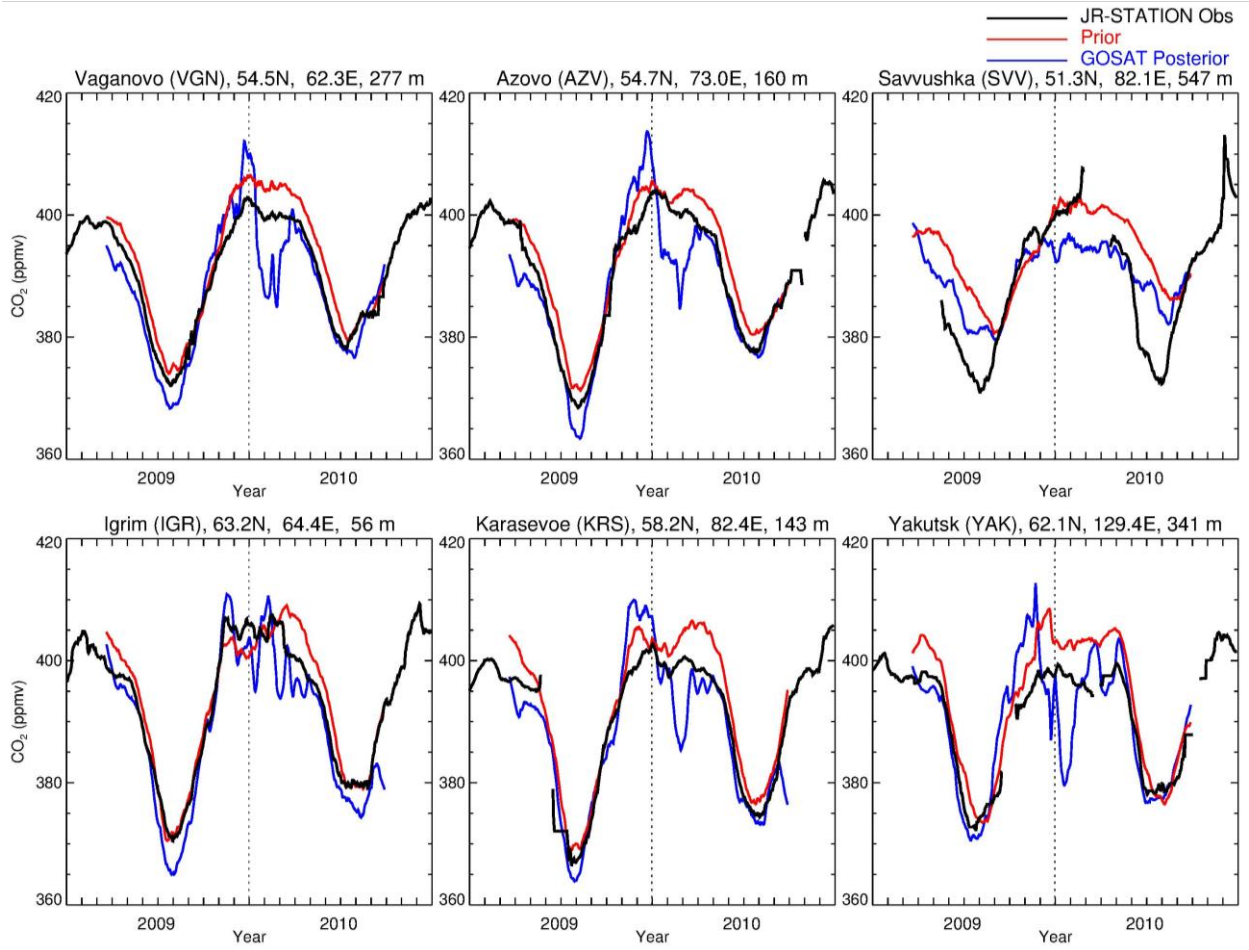


Figure 15~~7~~. Evaluation of the prior model and GOSAT-only inversion against JR-STATION in situ observations in Siberia. Shown are daily afternoon average (1200-1700 local time) mole fractions from the highest level on each tower, the time series of which are smoothed with a 31-day window. Sites are arranged from west to east, first at lower latitudes and then at higher latitudes, excluding those with data gaps in the summer. Elevations shown include intake heights on towers.

Table S1. In Situ Observation Sites Used in Inversions.

Site Code[a]	Name and Country	Latitude	Longitude	Elevation (m ASL)	Agency	Obs Type[b]	Mismatch (ppm)[c]
ABP	Arembepe, Bahia, Brazil	-12.77	-38.17	1	NOAA	F	1.25
ALT	Alert, Nunavut, Canada	82.45	-62.52	190	NOAA	F	0.75
AMT107	Argyle, Maine, U.S.	45.03	-68.68	53	NOAA	C	--
ASC	Ascension Island, UK	-7.97	-14.4	85	NOAA	F	0.4
ASK	Assekrem, Algeria	23.18	5.42	2710	NOAA	F	0.75
AZR	Terceira Island, Azores, Portugal	38.77	-27.38	19	NOAA	F	0.75
BAL	Baltic Sea, Poland	55.35	17.22	3	NOAA	F	4
BAO300	Boulder Atmospheric Observatory, Colorado, U.S.	40.05	-105	1584	NOAA	C	--
BKT	Bukit Kototabang, Indonesia	-0.2	100.32	845	NOAA	F	4
BMW	Tudor Hill, Bermuda, UK	32.27	-64.88	30	NOAA	F	0.75
BRW	Barrow, Alaska, U.S.	71.32	-156.6	11	NOAA	F	0.75
BSC	Black Sea, Constanta, Romania	44.17	28.67	0	NOAA	F	4
CBA	Cold Bay, Alaska, U.S.	55.2	-162.72	21	NOAA	F	0.75
CGO	Cape Grim, Tasmania, Australia	-40.68	144.69	94	NOAA	F	0.4
CHR	Christmas Island, Kiribati	1.7	-157.17	0	NOAA	F	0.4
CIB	Centro de Investigacion de la Baja Atmosfera, Spain	41.81	-4.93	845	NOAA	F	2.5
CPT	Cape Point, South Africa	-34.35	18.49	230	NOAA	F	0.75
CRZ	Crozet Island, France	-46.45	51.85	197	NOAA	F	0.4
DRP	Drake Passage	-59	-64.69	0	NOAA	F	0.4
DSI	Dongsha Island, Taiwan	20.7	116.73	3	NOAA	F	0.75
EIC	Easter Island, Chile	-27.15	-109.45	47	NOAA	F	0.75
GMI	Guam, Mariana Islands	13.43	144.78	0	NOAA	F	0.75
HBA	Halley Station, Antarctica, UK	-75.58	-26.21	30	NOAA	F	0.4
HPB	Hohenpeissenberg, Germany	47.8	11.01	936	NOAA	F	4

HSU	Humboldt State University, U.S.	41.06	-124.75	0	NOAA	F	0.75
HUN	Hegyhatsal, Hungary	46.95	16.65	248	NOAA	F	4
ICE	Storhofdi, Vestmannaeyjar, Iceland	63.4	-20.29	118	NOAA	F	0.75
IZO	Izana, Tenerife, Canary Islands, Spain	28.3	-16.48	2373	NOAA	F	0.75
KEY	Key Biscayne, Florida, U.S.	25.67	-80.2	1	NOAA	F	1.25
KUM	Cape Kumukahi, Hawaii, U.S.	19.52	-154.82	3	NOAA	F	0.75
KZD	Sary Taukum, Kazakhstan	44.45	77.57	595	NOAA	F	1.25
KZM	Plateau Assy, Kazakhstan	43.25	77.88	2519	NOAA	F	1.25
LEF396	Park Falls, Wisconsin, U.S.	45.93	-90.27	472	NOAA	C	--
LLB	Lac La Biche, Alberta, Canada	54.95	-112.45	540	NOAA	F	1.5
LLN	Lulin, Taiwan	23.47	120.87	2862	NOAA	F	1.25
LMP	Lampedusa, Italy	35.52	12.62	45	NOAA	F	0.75
MEX	High Altitude Global Climate Observation Center, Mexico	18.98	-97.31	4464	NOAA	F	1.25
MHD	Mace Head, County Galway, Ireland	53.33	-9.9	5	NOAA	F	1.25
MID	Sand Island, Midway, U.S.	28.21	-177.38	11	NOAA	F	0.75
MKN	Mt. Kenya, Kenya	-0.05	37.3	3644	NOAA	F	1.25
MLO	Mauna Loa, Hawaii, U.S.	19.54	-155.58	3397	NOAA	F	0.75
MNM	Minamitorishima, Japan	24.3	153.97	8	JMA	C	--
NAT	Farol De Mae Luiza Lighthouse, Brazil	-5.51	-35.26	50	NOAA	F	0.75
NMB	Gobabeb, Namibia	-23.58	15.03	456	NOAA	F	1.25
NWR	Niwot Ridge, Colorado, U.S.	40.05	-105.58	3523	NOAA	F	0.75
OXK	Ochsenkopf, Germany	50.07	11.8	1022	NOAA	F	1.25
PAL	Pallas-Sammaltunturi, GAW Station, Finland	67.97	24.12	565	NOAA	F	1.25
POCN00	Pacific Ocean (0 N)	0	-163	10	NOAA	F	0.4
POCN05	Pacific Ocean (5 N)	5	-158	10	NOAA	F	0.4
POCN10	Pacific Ocean (10 N)	10	-152	10	NOAA	F	0.4

POCN15	Pacific Ocean (15 N)	15	-147	10	NOAA	F	0.4
POCN20	Pacific Ocean (20 N)	20	-140	10	NOAA	F	0.4
POCN25	Pacific Ocean (25 N)	25	-134	10	NOAA	F	0.4
POCN30	Pacific Ocean (30 N)	30	-126	10	NOAA	F	0.4
POCS05	Pacific Ocean (5 S)	-5	-168	10	NOAA	F	0.4
POCS10	Pacific Ocean (10 S)	-10	-174	10	NOAA	F	0.4
POCS15	Pacific Ocean (15 S)	-15	-178	10	NOAA	F	0.4
POCS20	Pacific Ocean (20 S)	-20	-178.5	10	NOAA	F	0.4
POCS25	Pacific Ocean (25 S)	-25	174	10	NOAA	F	0.4
POCS30	Pacific Ocean (30 S)	-30	169	10	NOAA	F	0.4
PSA	Palmer Station, Antarctica, U.S.	-64.92	-64	10	NOAA	F	0.4
PTA	Point Arena, California, U.S.	38.95	-126.23	17	NOAA	F	2.5
RPB	Ragged Point, Barbados	13.17	-59.43	15	NOAA	F	0.75
RYO	Ryori, Japan	39.03	141.83	260	JMA	C	--
SCT305	Beech Island, South Carolina, U.S.	33.41	-81.83	115	NOAA	C	--
SDZ	Shangdianzi, China	40.65	117.12	293	NOAA	F	4
SEY	Mahe Island, Seychelles	-4.67	55.17	2	NOAA	F	0.4
SGP	Southern Great Plains, Oklahoma, U.S.	36.8	-97.5	314	NOAA	F	1.25
SHM	Shemya Island, Alaska, U.S.	52.72	174.1	23	NOAA	F	1.25
SMO	Tutuila, American Samoa	-14.24	-170.57	42	NOAA	F	0.4
SPO	South Pole, Antarctica, U.S.	-89.98	-24.8	2810	NOAA	F	0.4
STM	Ocean Station M, Norway	66	2	0	NOAA	F	0.75
SUM	Summit, Greenland	72.58	-38.48	3210	NOAA	F	0.75
SYO	Syowa Station, Antarctica, Japan	-69.01	39.58	14	NOAA	F	0.4
TAP	Tae-ahn Peninsula, Republic of Korea	36.73	126.13	16	NOAA	F	4
TDF	Tierra del Fuego, Argentina	-54.87	-68.48	20	NOAA	F	0.4
THD	Trinidad Head, California, U.S.	41.05	-124.15	107	NOAA	F	1.25
UTA	Wendover, Utah, U.S.	39.9	-113.72	1327	NOAA	F	1.25
UUM	Ulaan Uul, Mongolia	44.45	111.1	1007	NOAA	F	1.25

WBI379	West Branch, Iowa, U.S.	41.72	-91.35	242	NOAA	C	--
WGC483	Walnut Grove, California, U.S.	38.27	-121.49	0	NOAA	C	--
WIS	Weizmann Institute of Science at the Arava Institute, Ketura, Israel	31.13	34.88	151	NOAA	F	1.25
WKT457	Moody, Texas, U.S.	31.32	-97.33	251	NOAA	C	--
WLG	Mt. Waliguan, China	36.29	100.9	3810	NOAA	F	1.25
YON	Yonagunijima, Japan	24.47	123.02	30	JMA	C	--
ZEP	Ny-Alesund, Svalbard, Norway and Sweden	78.91	11.88	474	NOAA	F	0.75

[a] Tower intake height appended where relevant

[b] F = Flask, C = Continuous

[c] Model-data mismatch component of observation error

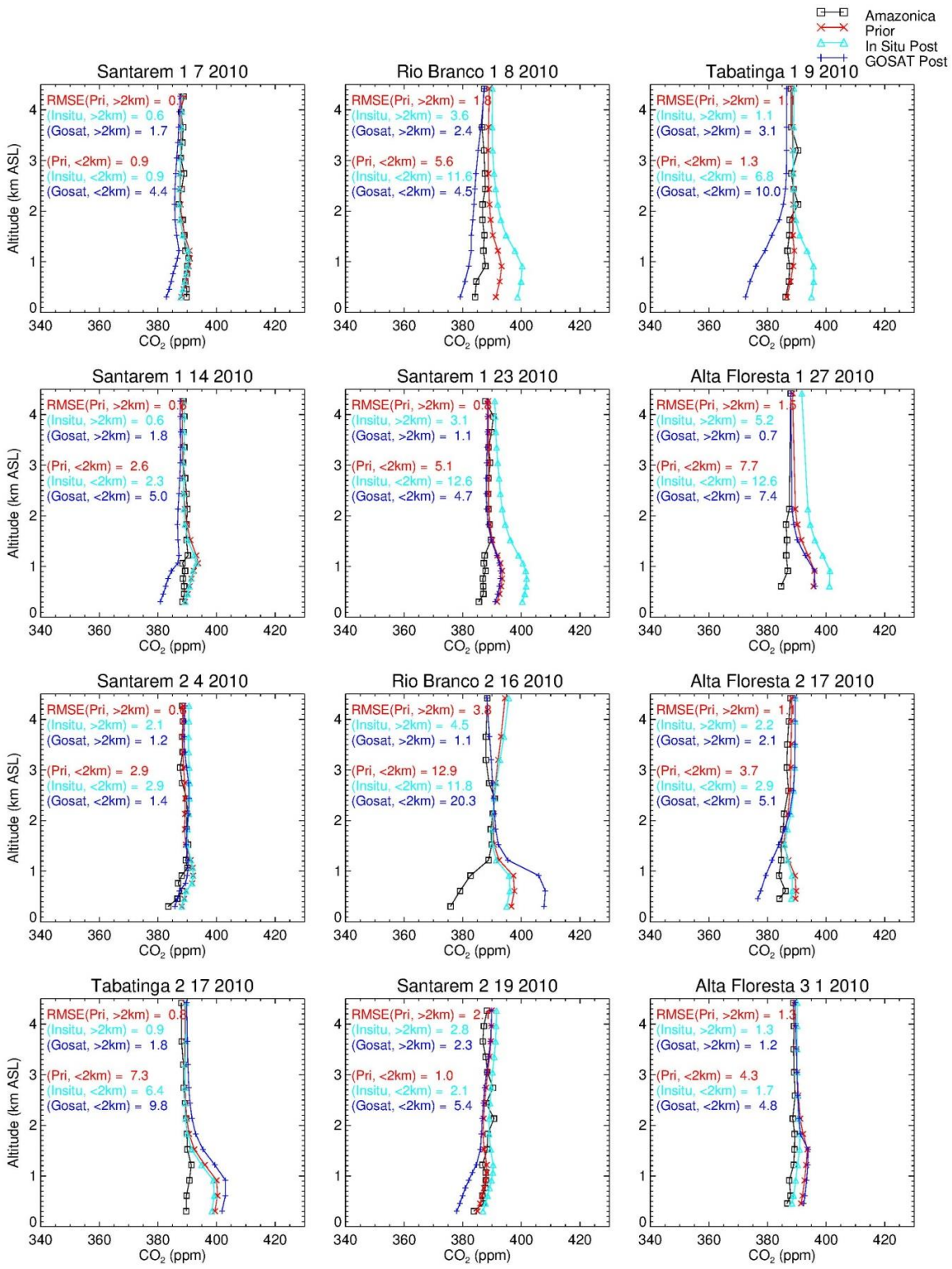


Figure S1. Comparison of prior, in situ-only posterior, GOSAT-only posterior, and Amazonica aircraft vertical profiles over 4 sites on different dates.

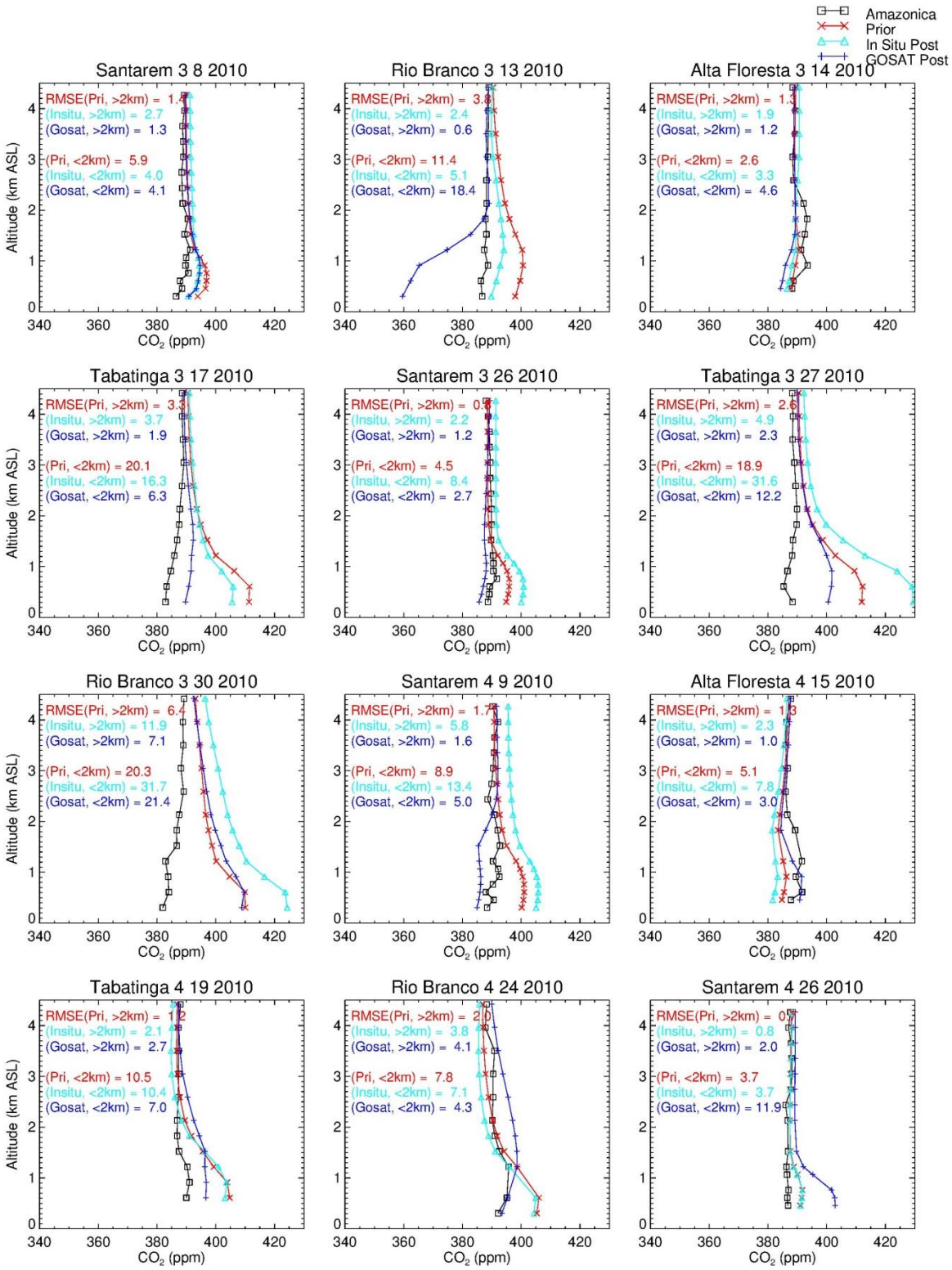


Figure S1. (continued)

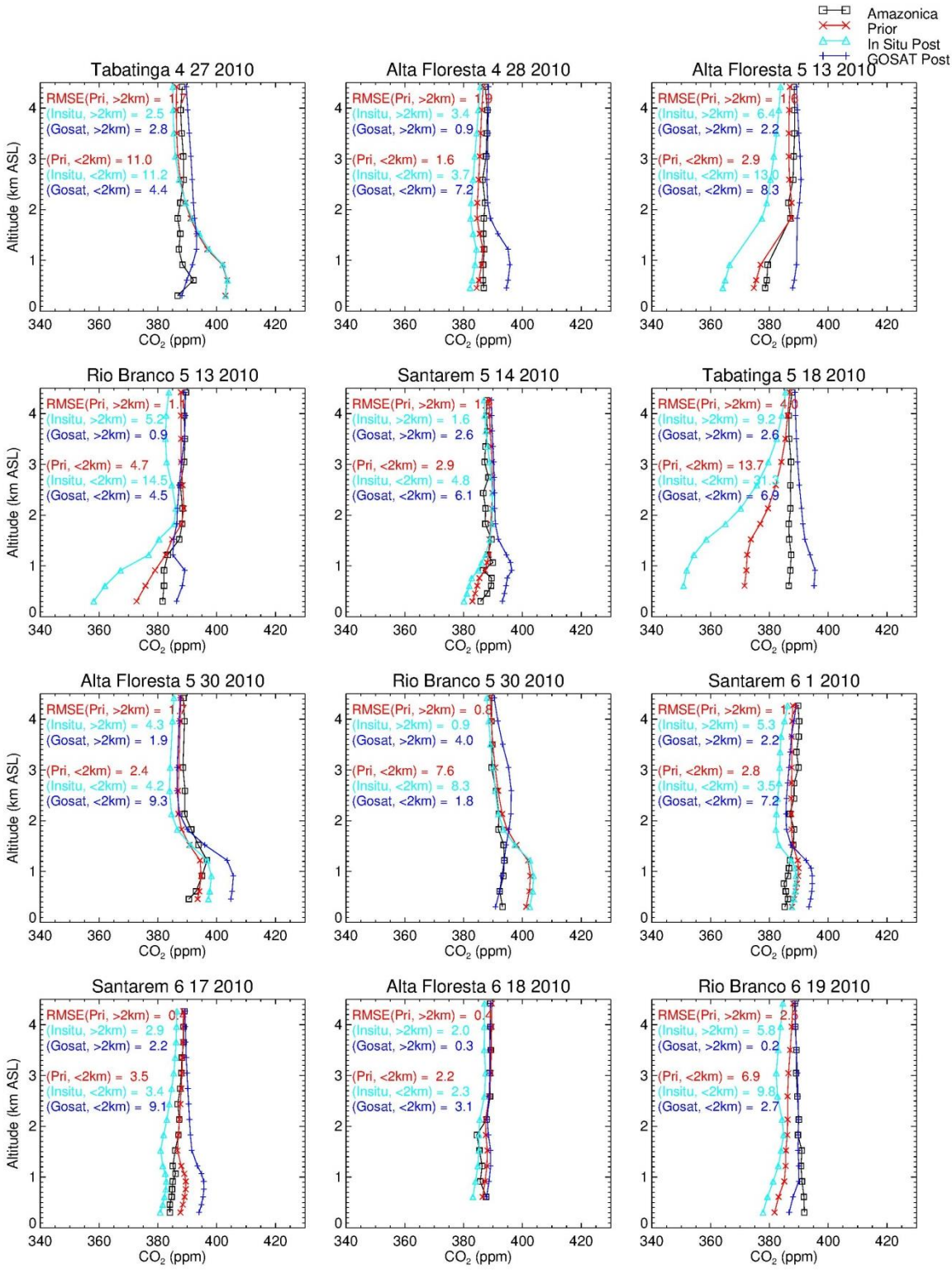


Figure S1. (continued)

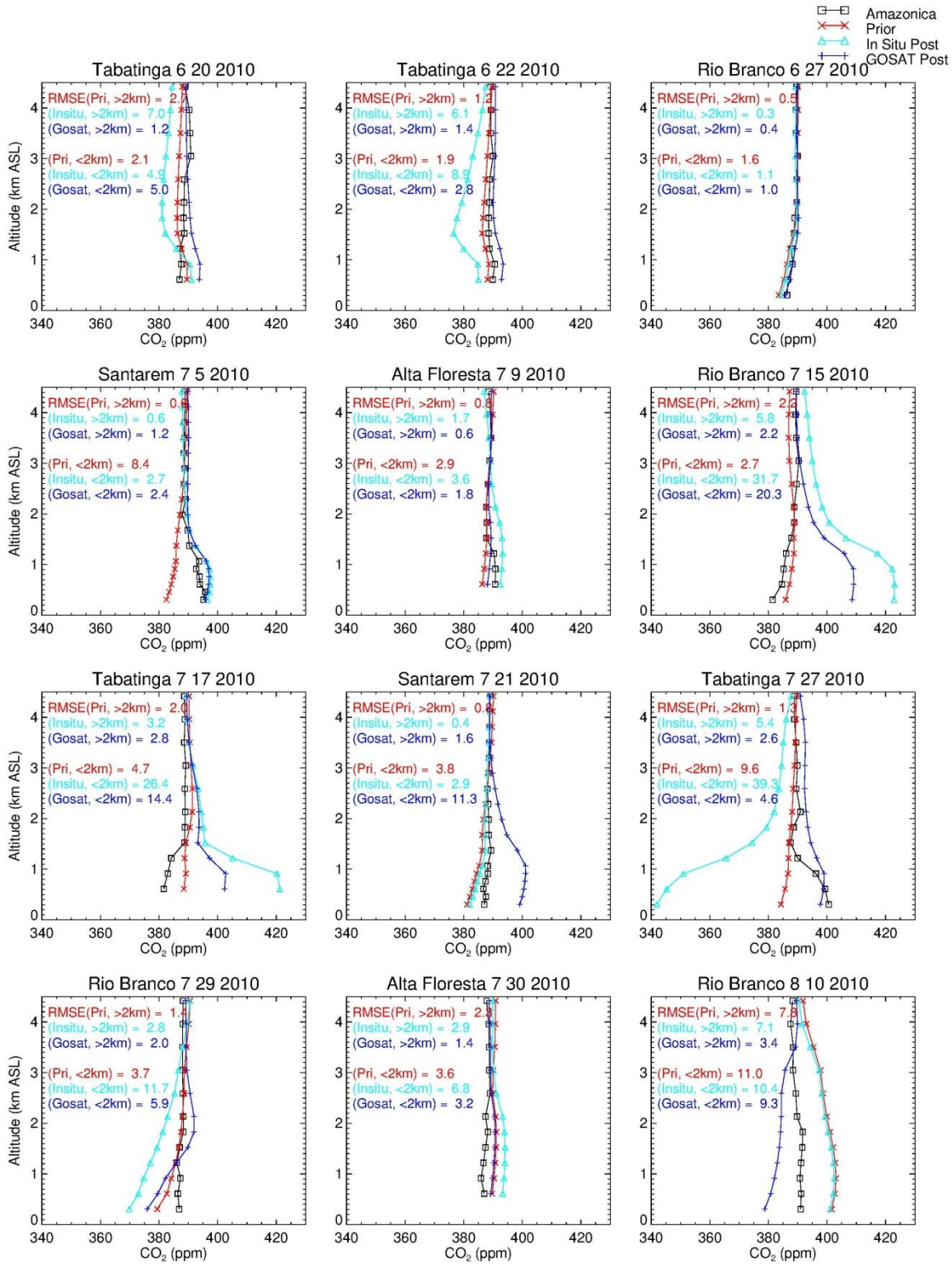


Figure S1. (continued)

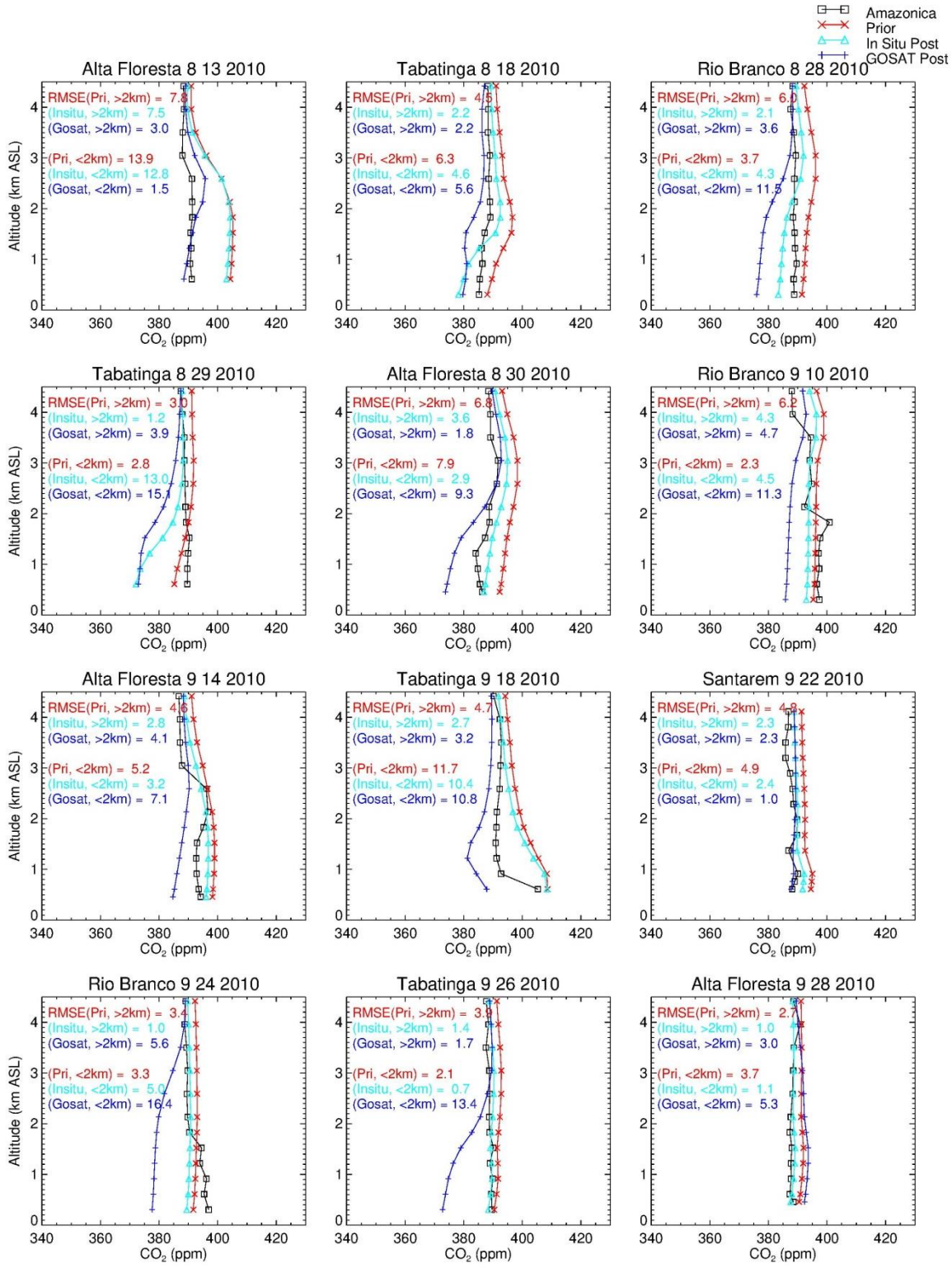


Figure S1. (continued)

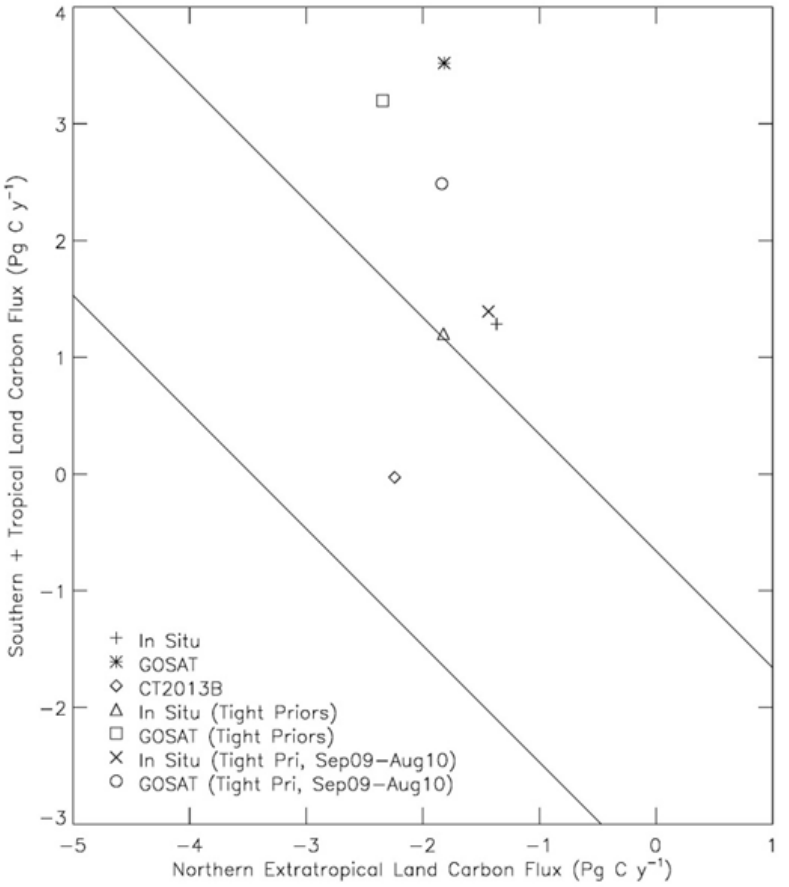


Figure S2. Posterior north-south land flux partitioning after Schimel et al. (2015). The diagonal lines are based on the global land carbon exchange (= land-use change emissions – land sink) estimated by GCP (2015) for the years relevant to the present analysis, i.e. 2009 and 2010, ± 1σ. Fluxes are for June 2009-May 2010 except where specified in the legend (for September 2009-August 2010). CT2013B refers to the CarbonTracker data assimilation system.

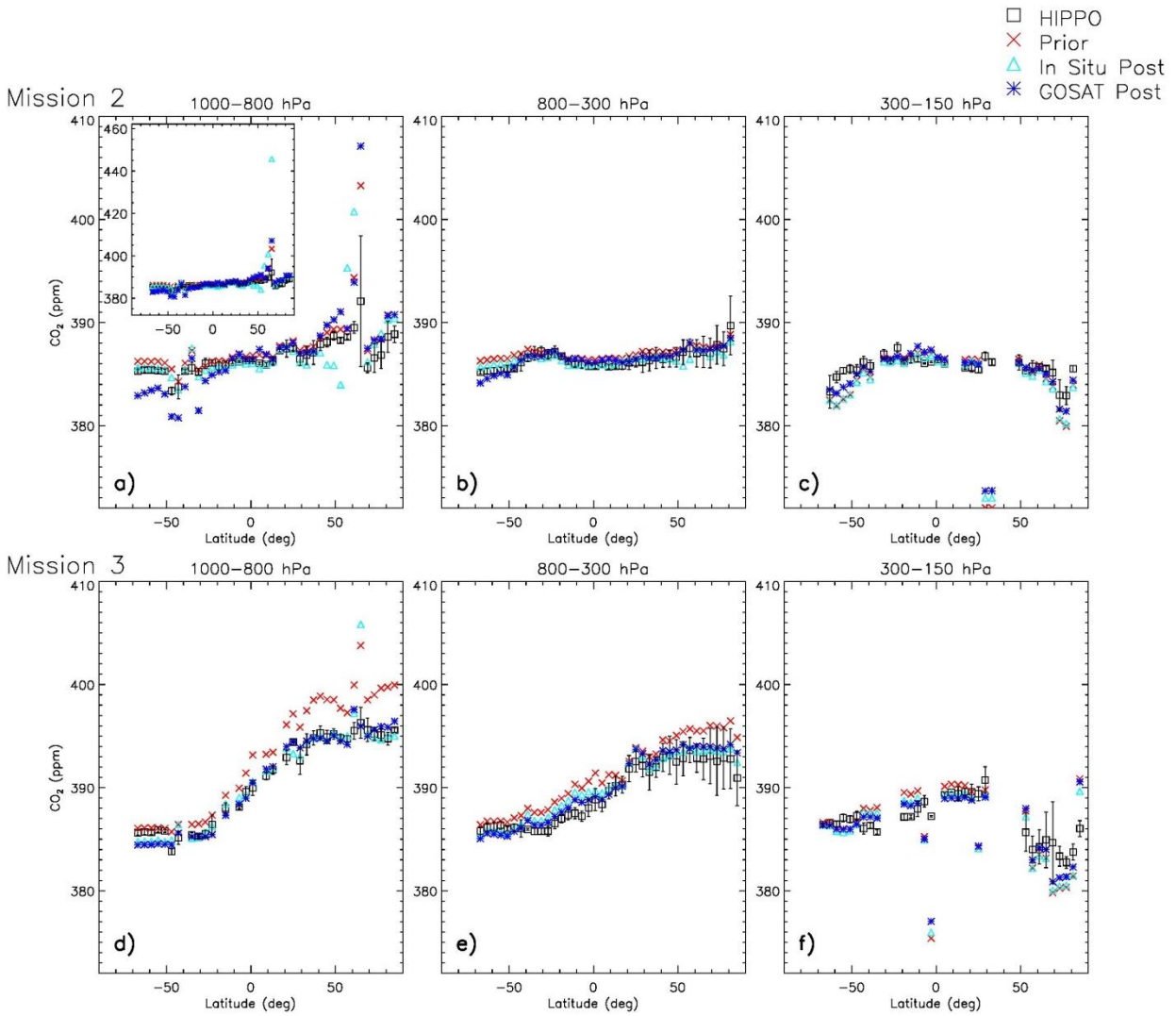


Figure S3. Same as Fig. 10 except showing inversions with tighter prior uncertainties.

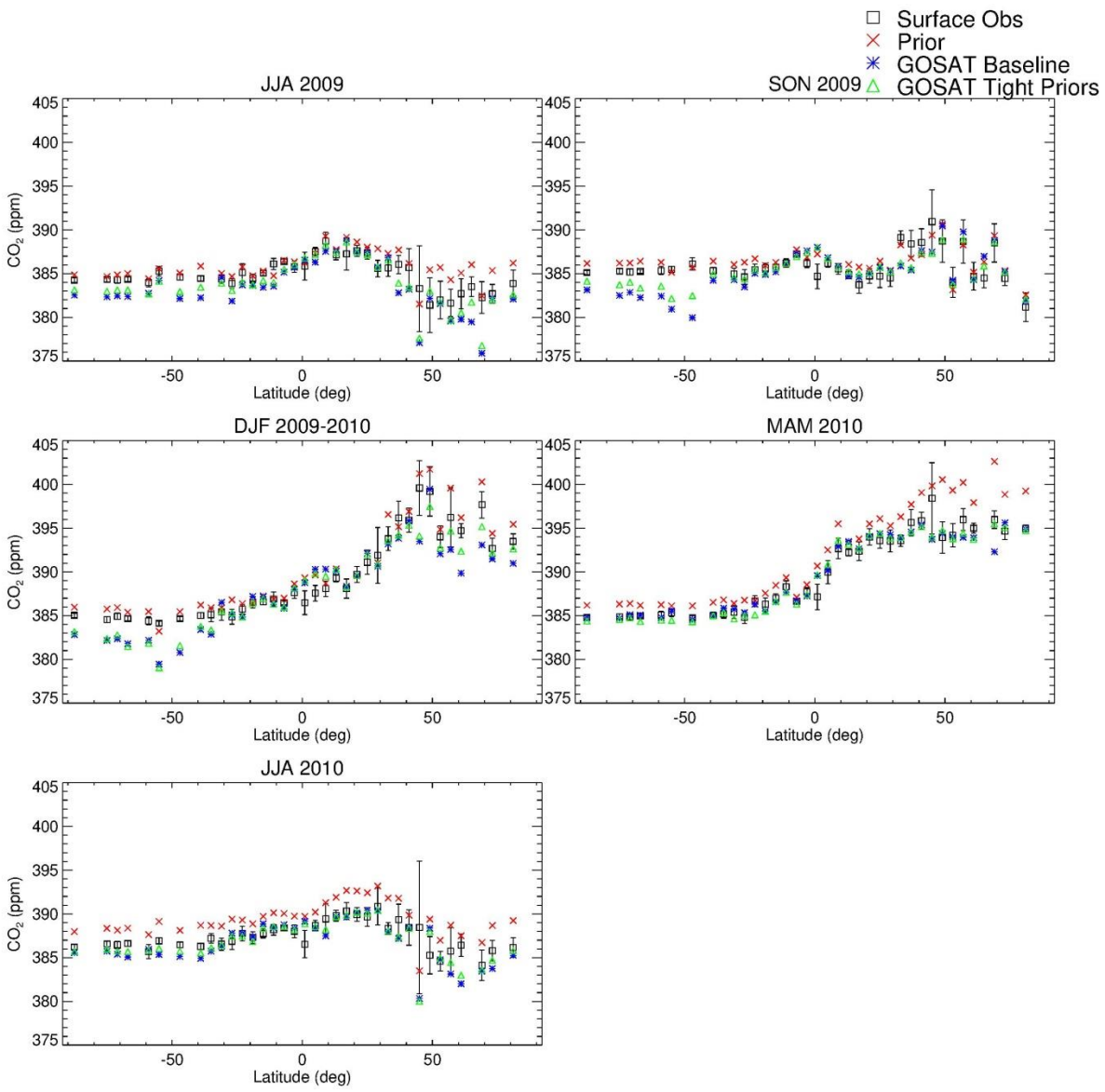


Figure S4. Same as Fig. 9 except showing GOSAT-only inversions with baseline vs. tighter prior uncertainties.