Improved constraints on northern extratropical CO2 fluxes obtained by combining surface-based and space-based atmospheric CO2 measurements

Brendan Keith Aidan Byrne[1,1], Junjie Liu[2,2], Meemong Lee[2,2], Ian T. Baker[3,3], Kevin W. Bowman[4,4], Nicholas Michael Deutscher[5,5], dietrich feist[6], David W.T Griffith[5,5], Laura T Iraci[7,7], Matthaus Kiel[2,2], John Kimball[8,8], Charles E. Miller[2,2], Isamu Morino[9,9], Nicholas Cody Parazoo[2,2], Christof Petri[10,10], Coleen Marie Roehl[11,11], Mahesh Sha[12,12], Kimberly Strong[13,13], Voltaire Almario Velazco[5,5], Paul O. Wennberg[11,11], Debra Wunch[13,13], and Dietrich G. Feist[14]

[1]Jet Propulsion Lab
[2]Jet Propulsion Laboratory
[3]Colorado State University
[4]Jet Propulsion Lab (NASA)
[5]University of Wollongong
[6]Max Planck Institute for Biogeochemistry
[7]NASA Ames Research Center
[8]University of Montana
[9]National Institute for Environmental Studies
[10]University of Bremen
[11]California Institute of Technology
[12]Royal Belgian Institute for Space Aeronomy
[13]University of Toronto
[14]Ludwig-Maximilians-Universität München

November 30, 2022

Abstract

Top-down estimates of CO2 fluxes are typically constrained by either surface-based or space-based CO2 observations. Both of these measurement types have spatial and temporal gaps in observational coverage that can lead to biases in inferred fluxes. Assimilating both surface-based and space-based measurements concurrently in a flux inversion framework improves observational coverage and reduces sampling biases. This study examines the consistency of flux constraints provided by these different observations and the potential to combine them by performing a series of six-year (2010–2015) CO2 flux inversions. Flux inversions are performed assimilating surface-based measurements from the in situ and flask network, measurements from the Total Carbon Column Observing Network (TCCON), and space-based measurements from the Greenhouse Gases Observing Satellite (GOSAT), or all three datasets combined. Combining the datasets results in more precise flux estimates for sub-continental regions relative to any of the datasets alone. Combining the datasets also improves the accuracy of the posterior fluxes, based on reduced root-mean-square differences between posterior-flux-simulated CO2 and aircraft-based CO2 over midlatitude regions (0.35–0.50˜ppm) in comparison to GOSAT (0.39–0.57˜ppm), TCCON (0.52–0.63˜ppm), or in situ and flask measurements (0.45–0.53˜ppm) alone. These results suggest that surface-based and GOSAT measurements give complementary constraints on CO2 fluxes in the northern extratropics and can be combined in flux inversions to improve observational coverage. This stands in contrast with many earlier attempts to combine these datasets and suggests that improvements in the NASA Atmospheric CO2 Observations from Space (ACOS) retrieval algorithm have significantly improved the consistency of space-based and surface-based flux constraints.

Improved constraints on northern extratropical CO_2 fluxes obtained by combining surface-based and space-based atmospheric CO_2 measurements

B. Byrne[1], J. Liu[2], M. Lee[2], I. Baker[3], K. W. Bowman[2], N. M. Deutscher[4], D. G. Feist[5], D. W. T. Griffith[4], L. T. Iraci[6], M. Kiel[2], J. S. Kimball[7], C. E. Miller[2], I. Morino[8], N. C. Parazoo[2], C. Petri[9], C. M. Roehl[10], M. K. Sha[11], K. Strong[12], V. A. Velazco[4], P. O. Wennberg[10,13], D. Wunch[12]

[1]NASA Postdoctoral Program Fellow, Jet Propulsion Laboratory, California Institute of Technology, CA, USA
[2]Jet Propulsion Laboratory, California Institute of Technology, CA, USA
[3]Atmospheric Science Department, Colorado State University, Fort Collins, CO, USA
[4]Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, Australia
[5]Max Planck Institute for Biogeochemistry, Jena, Germany
[6]Atmospheric Science Branch, NASA Ames Research Center, Moffett Field, CA 94035, USA
[7]Numerical Terradynamic Simulation Group, W.A. Franke College of Forestry & Conservation, The University of Montana, Missoula, MT 59812, USA
[8]Satellite Observation Center, Center for Global Environmental Research, National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan
[9]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[10]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[11]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium
[12]Department of Physics, University of Toronto, Toronto, Ontario, Canada
[13]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, USA

Key Points:

- Consistent flux constraints provided by surface in situ and flask, TCCON, and GOSAT measurements of atmospheric CO_2.
- Combining data sets improves agreement between modeled and measured aircraft-based CO_2 measurements.
- Improvements in NASA ACOS retrieval explain improved consistency of space-based and surface-based CO_2.

Corresponding author: Brendan Byrne, `brendan.k.byrne@jpl.nasa.gov`

Abstract

Top-down estimates of CO_2 fluxes are typically constrained by either surface-based or space-based CO_2 observations. Both of these measurement types have spatial and temporal gaps in observational coverage that can lead to biases in inferred fluxes. Assimilating both surface-based and space-based measurements concurrently in a flux inversion framework improves observational coverage and reduces sampling biases. This study examines the consistency of flux constraints provided by these different observations and the potential to combine them by performing a series of six-year (2010–2015) CO_2 flux inversions. Flux inversions are performed assimilating surface-based measurements from the in situ and flask network, measurements from the Total Carbon Column Observing Network (TCCON), and space-based measurements from the Greenhouse Gases Observing Satellite (GOSAT), or all three datasets combined. Combining the datasets results in more precise flux estimates for sub-continental regions relative to any of the datasets alone. Combining the datasets also improves the accuracy of the posterior fluxes, based on reduced root-mean-square differences between posterior-flux-simulated CO_2 and aircraft-based CO_2 over midlatitude regions (0.35–0.50 ppm) in comparison to GOSAT (0.39–0.57 ppm), TCCON (0.52–0.63 ppm), or in situ and flask measurements (0.45–0.53 ppm) alone. These results suggest that surface-based and GOSAT measurements give complementary constraints on CO_2 fluxes in the northern extratropics and can be combined in flux inversions to improve observational coverage. This stands in contrast with many earlier attempts to combine these datasets and suggests that improvements in the NASA Atmospheric CO_2 Observations from Space (ACOS) retrieval algorithm have significantly improved the consistency of space-based and surface-based flux constraints.

1 Introduction

Observations of atmospheric CO_2 provide a constraint on the net surface–atmosphere CO_2 flux, and are critical for monitoring carbon flux changes. This has motivated observational programs that measure atmospheric CO_2, including a global network of surface-based in situ and flask monitoring sites, the Total Carbon Column Observing Network (TCCON) of ground-based spectrometers (Wunch et al., 2011) and several satellite missions (Crisp et al., 2004; Yokota et al., 2009). These observations have provided many insights into the terrestrial carbon cycle (Keeling, 1960; Bolin & Keeling, 1963; Bacastow, 1976; Tans et al., 1989; Keeling et al., 1996; Bowman et al., 2017; J. Liu et al., 2017; Chatterjee et al., 2017). However, current measurement programs are unable to continuously monitor CO_2 with global coverage, resulting in observational gaps. These spatial and temporal gaps in observations of atmospheric CO_2 can introduce artifacts into NEE estimates, leading to difficulties in constraining carbon fluxes on regional scales (J. Liu et al., 2014; Byrne et al., 2017; Basu et al., 2018).

Different observing systems have different gaps in the observational coverage. Space-based measurements retrieve atmospheric CO_2 from measurements of reflected sunlight. This results in highly seasonal observational coverage in extratropical regions. Seasonal differences in observational coverage are further exasperated by challenging retrievals over snow (Nassar et al., 2014), and seasonal variations in cloud cover. In contrast, surface-based measurements of atmospheric CO_2 typically have comparatively uniform temporal coverage, but poor spatial coverage. Surface measurements sites most densely cover the northern extratropics (particularly North America and Europe) but have sparse coverage elsewhere (Byrne et al., 2017).

In the northern extratropics, surface-based and space-based atmospheric CO_2 measurements provide complementary observational coverage in space and time, respectively. Yet, few studies have attempted to combine surface-based and space-based atmospheric CO_2 measurements to obtain top down constraints on fluxes across the northern latitudes. Chevallier et al. (2011) found consistency between the surface-air-sample-based and the TCCON-based inversions, suggesting that flux inversions combining both data sources could be performed. Houweling et al. (2015) performed a series of CO_2 flux inversions assimilating measurements from the Greenhouse Gases Observing Satellite (GOSAT) and surface-based CO_2 measurements. They found that comparisons between posterior CO_2 fields and aircraft data did not show significant differences between inversions assimilating surface-based or space-based measurements, and that the largest differences were driven by the inversion set up. However, they also found that the two datasets gave large differences in the spatial distribution of the CO_2 sink, with GOSAT flux inversions having increased uptake in the northern extratropics by $\sim$1 PgC. When both datasets were combined, they found that the posterior fluxes did not recover the observed meridional gradient in CO_2 (which was also found for the GOSAT flux inversions), suggesting that the biases in retrieved GOSAT X_{CO_2} could be adversely impacting the results. Another study, Wang et al. (2018), assimilated both GOSAT measurements and surface-based atmospheric CO_2 measurements in a batch Bayesian synthesis inversion. They found that the differences in observational coverage of the ground-based and space-based datasets were complementary, resulting in smaller posterior uncertainty estimates when both datasets are assimilated than either dataset alone. Similarly, in a set of regional Observing System Simulation Experiments (OSSEs), Fischer et al. (2017) showed reduced uncertainty in biosphere and fossil fuel emissions in California by combining space-based X_{CO_2} and surface-based flask and in situ measurements.

In this study, we further investigate combining ground-based and space-based measurements of atmospheric CO_2 to provide estimates of NEE globaly, but we focus on northern extra-tropical regions were surface-based and aircraft-based measurements are most densly concentrated.. We perform a series of six-year flux inversions (2010–2015, inclusive) assimilating surface-based measurements from the in situ and flask measurement network, TCCON column-averaged dry-air CO_2 mole fractions (X_{CO_2}), GOSAT X_{CO_2} measurements, and all three datasets combined. For each set of measurements, we perform three flux inversions applying different prior NEE flux and error constraints. From the spread in posterior fluxes due to prior constraints, we quantify the precision to which these datasets constrain posterior fluxes. Spatial structures in the posterior fluxes are examined through comparisons between posterior-NEE-simulated X_{CO_2} and Orbiting Carbon Observatory 2 (OCO-2) X_{CO_2} measurements and the accuracy of posterior-NEE-simulated CO_2 is examined through comparisons with aircraft-based CO_2 measurements.

The paper is outlined as follows. Section 2 describes the measurements used in this study and Sec. 3 describes the flux inversion set-up. The posterior CO_2 fields obtained by the flux inversions are compared with OCO-2 and aircraft-based measurements in Sec. 4.1. We then examine the six-year-mean seasonal cycle and annual net fluxes (Sec. 4.2) and interannual variability (Sec. 4.3) obtained by the flux inversions. Finally, the implications of the results are discussed in Sec. 5 and conclusions are given in Sec. 6.

2 Data

2.1 Surface-based in situ and flask measurements

Surface-based measurements of boundary layer atmospheric CO_2 can be performed using an in situ gas analyzer or by taking a flask sample, which is then returned to a lab and analyzed. A number of different groups from around the world collect surface CO_2 observations. We assimilate measurements from version 4.1 of the GLOBALVIEW plus package (Masarie et al., 2014; Cooperative Global Atmospheric Data Integration Project, 2018) and the Japan-Russia Siberian Tall Tower Inland Observation Network (JR-STATION) of nine tower sites in Siberia (Sasakawa et al., 2010, 2013).

The GLOBALVIEW v4.1 package incorporates data from many observing sites around the world and is specifically prepared for use in data assimilation studies. We include measurements from the Integrated Carbon Observation System (ICOS RI, 2019) in our analysis. We assimilate GLOBALVIEW v4.1 measurements from surface in situ and flask sites, tower sites, and ship-based measurements. Data is only assimilated if the measurements are assimilated by NOAA's CarbonTracker, version CT2017 (CT_assim = 0). Measurements are assimilated at the intake height above the model surface over land, and at the intake height above sea level for ocean grid cells. For surface-based flask and in situ measurements, most of the measurement error applied for assimilation is due to representativeness errors (inability to model these measurements). We use the model-data-mismatch (mdm) as the measurement errors. This is the error value placed on each measurement in the assimilation system, and is meant to express the statistics of simulated-minus-observed CO2 residuals expected if CarbonTracker were using perfect surface fluxes.

JR-STATION is a network of nine towers (http://www.cger.nies.go.jp/en/climate/pj1/tower/). On these towers, high inlet measurements are obtained over the 17–20th minutes of each hour and the low inlet data is obtained from 37–40th minutes of each hour, these 3-minute averages are the taken to be representative of the hourly means for each inlet. We filter the measurements by removing all measurements where the vertical gradient in CO_2 exceeds 0.5 ppm (to remove measurements when the boundary layer is not well-mixed), and use the measured value at the highest intake for the measurement. For each site the errors (in ppm) are prescribed to be constant throughout a given month, the errors are the errors range from 3 ppm in winter to 7 ppm in summer, to account for both measurement and representativeness errors. These error estimates were chosen because they are comparable to the error estimates for tower sites in the GLOBALVIEW plus v4.1 package.

We remove outliers and poorly modeled measurements by filtering out measurements for which the difference between the prior-NEE-simulated measurements and actual measurements exceeds three standard deviations of the measurement uncertainty (See Sec. 3 for details on the forward model simulations). We also remove measurements for which the difference between prior simulated CO_2 and measurement exceeds 10 ppm, as these are assumed to be poorly simulated by the model. This filtering removes $\sim$8% of the measurements. For each site, the data is only assimilated between 11 a.m. and 4 p.m. local time.

2.2 Aircraft-based measurements

Aircraft measurements are used for the evaluation of posterior atmospheric CO_2 fields. Aircraft data are obtained from the version 4.1 of the GLOBALVIEW plus dataset. Comparisons between measured and modeled atmospheric CO_2 are performed over three distinct regions: East Asia, North America, and Alaska/Arctic (Fig. S1). Aircraft measurements over East Asia come exclusively from the Comprehensive Observation Network for Trace gases by Airliner (CONTRAIL) program (Machida et al., 2008, 2018). Aircraft data over Alaska/Arctic and North America originate from the NOAA Global Greenhouse Gas Reference Network's aircraft program (Sweeney et al., 2015) and HIAPER Pole-to-Pole Observations (HIPPO) (Wofsy, 2011). The number of hourly-mean measurements per month between 3–8 km in altitude above sea level (asl) are shown in Fig. S2.

2.3 TCCON measurements

TCCON is a network of ground-based Fourier transform spectrometers that record solar absorption spectra in the near-infrared from which, among other gases, X_{CO_2} is estimated (Wunch et al., 2011). CO_2 abundances are retrieved using a non-linear least squares approach from absorption lines in the near-infrared spectral region. The column-averaged dry-air mole fractions of CO_2 (X_{CO_2}) is calculated by taking the ratio of the column abundance of CO_2 to O_2 (scaled by the mean O_2 concentration), resulting in high precision

Table 1. TCCON sites used in this study.

Site Name	Lat	Lon	Start Date	Reference
Eureka	80.05 N	86.42 W	25 Jul 2010	Strong et al. (2017)
Orleans	47.97 N	2.11 E	29 Aug 2009	Warneke et al. (2017)
Park Falls	45.95 N	90.27 W	02 Jun 2004	Wennberg, Roehl, et al. (2017)
Rikubetsu	43.46 N	143.77 E	16 Nov 2013	Morino et al. (2017)
Lamont	36.60 N	97.49 W	06 Jul 2008	Wennberg, Wunch, et al. (2017)
Edwards	34.96 N	117.88 W	20 Jul 2013	Iraci et al. (2017)
Ascension Island	7.92 S	14.33 W	22 May 2012	Feist et al. (2017)
Darwin	12.46 S	130.93 E	28 Aug 2005	Griffith, Deutscher, et al. (2017)
Reunion Island	20.90 S	55.49 E	16 Sep 2011	De Mazière et al. (2017)
Wollongong	34.41 S	150.88 E	26 Jun 2008	Griffith, Velazco, et al. (2017)

(<0.25% in CO_2) X_{CO_2} measurements. The TCCON strives to achieve the best site-to-site precision and accuracy possible. Systematic biases that are consistent throughout the network are fully accounted for by scaling the TCCON retrieval results to the WMO scale via aircraft and AirCore profiles (Wunch et al., 2010). Moreover, the TCCON sets guidelines to ensure that the instrumentation at each site is as similar as possible, and that the retrieval software, including the spectroscopic line lists and line shapes, is identical for each site. However, site-specific differences (e.g. instrumental line shape) can cause residual site-to-site biases (Wunch et al., 2010) which might introduce biases in flux inversions.

For this study, TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA [https://tccondata.org]. We include data from TCCON sites that have mean biases of less than 0.5 ppm relative to both the OCO-2 target-mode X_{CO_2} and the posterior-simulated X_{CO_2} from the surface-only flux inversions. The sites included in this study, which provide data during the years 2010–2015, are given in Table 1. Sites that are excluded from this study are excluded due to several factors that cause apparent biases to be greater than 0.5 ppm. These factors include: proximity to large CO_2 sources (e.g., cities), proximity to large topographic variability, and in a few cases, known TCCON instrument biases for which a solution either has been applied, or will be applied in an upcoming TCCON data version. Note that the threshold of 0.5 ppm is somewhat arbitrary. This value was set because most sites outside of this threshold are in heavily observed regions (e.g., Europe), which are expected to be well constrained by other datasets (Byrne et al., 2017), or in the Southern Hemisphere and not expected to have a large impact on the performance of the flux inversions in the northern mid-latitudes.

In this study, the TCCON data are filtered to remove measurements with solar zenith angles greater than 70 degrees. Measurements are then binned into hourly medians for each site. Only hours with five or more measurements are included. Measurements are only assimilated between 11am-3pm local time for the flux inversions, to minimize potential biases relating to errors in the prescribed diurnal cycle of NEE.

2.4 Space-based measurements

We assimilate X_{CO_2} measured by the Thermal And Near-infrared Sensor for carbon Observations Fourier Transform Spectrometer (TANSO-FTS) aboard GOSAT. GOSAT was launched in February 2009 in a sun-synchronous orbit, with a repeat cycle of 3 days that produces 44 separate ground track repeats (Yoshida et al., 2013). The footprint of the GOSAT measurements has a diameter of about 10 km. Since August 2010, TANSO-FTS has been measuring with a 3-point cross-track pattern with 263 km cross track separation, resulting in a swath of 526 km. Measurements have an along-track separation of 283 km (Crisp et al., 2012). We use version 7.3 of the NASA Atmospheric CO_2 Observations from Space (ACOS) GOSAT measurements in this analysis. A detailed decription of ACOS retrieval algorithm is available in O'Dell et al. (2012) and Crisp et al. (2012), with recent updates described in Eldering et al. (2017) and O'Dell et al. (2018). We assimilate all high gain (H-Gain) nadir measurements from the TANSO-FTS shortwave infrared (SWIR) band that pass the quality flag requirement.

Measurements from OCO-2 are used for comparisons with the posterior CO_2 fields. OCO-2, launched in July 2014, is a space-based spectrometer in a Sun-synchronous orbit that measures reflected solar radiation to infer X_{CO_2} with a footprint of about 3 km^2. It has a repeat cycle of 16 days, resulting in 233 separate ground track repeats. OCO-2 has a swath of 10 km and collects eight adjacent, spatially resolved samples every 0.333 s, resulting in roughly 24 soundings per second. We downloaded version 9 of the ACOS OCO-2 lite files from the CO_2 Virtual Science Data Environment (https://co2.jpl.nasa.gov/). Measurements are averaged into super-obs at $1° \times 1°$ resolution grids following J. Liu et al. (2017), with the additional requirement that there must be a minimum of eight OCO-2 observations within each $1°\times1°$ gridbox. We combine land nadir and land glint measurements for the analysis.

3 Flux inversions

Flux inversions are performed with the Greenhouse Gas Framework – Flux (GHGF-Flux) inversion system. GHGF-Flux is a flux inversion system developed under the NASA's Carbon Monitoring System (CMS) project. The GHGF is capable of jointly assimilating multi-platform observations of CH_4, CO, CO_2, and OCS. The GHGF inherits the chemistry transport model from the GEOS-Chem and the adjoint analysis methods from the GEOS-Chem-adjoint.

Chemical transport is driven by the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) meteorology produced with version 5.12.4 of the GEOS atmospheric data assimilation system (Gelaro et al., 2017). To perform tracer transport, these fields are regridded to $4°\times 5°$ horizontal resolution and archived with a temporal resolution of 6 h except for surface quantities and mixing depths, which have a temporal resolution of 3 h. Tracer transport is performed at 30 min time steps.

For all inversions, we optimize 14 day scaling factors for daily net NEE and ocean fluxes, except for the final temporal grouping of each year, which is padded with 1–2 days so that the groupings cover the same day-of-year increments for each year. We use an assimilation window of approximately 18 months (October 7 to April 1 two years later) and keep posterior fluxes for one year (Jan 1 to Dec 31) then shift the inversion window forward one year. Using this method, we optimize NEE spanning 2010–2015. Initial conditions are generated by performing a two year inversion of surface in situ and flask measurements spanning 1 Jan 2008 to 31 Dec 2009. The stratosphere is then adjusted to match the zonal mean structure of Diallo et al. (2017) for October 2009 (adjusted by a few parts per million).

Prior NEE fluxes and errors differ between inversions, and are generated from three different models: the Simple Biosphere model (SiB3), the Carnegie-Ames-Stanford Approach model (CASA) and FLUXCOM. The motivation for using three different priors is that the posterior flux estimates may be sensitive to prior fluxes (Philip et al., 2019), thus using an ensemble of prior flux estimates provides an estimate of the precision to which the observations constrain fluxes. For all prior fluxes the annual total net flux has been adjusted to 4.6 PgC yr^{-1}, to match the mean atmospheric CO_2 growth rate. Details on the modeled NEE fluxes and prior errors are given in Appendix 7. The diurnal cycle in NEE is prescribed using the modeled diurnal cycle from SiB3 for the SiB3 flux inversions and the diurnal cycle from CASA for the CASA and FLUXCOM inversions.

Sensitivity tests found that the flux inversions were not sensitive to the prescribed diurnal NEE cycle. The ECCO-Darwin-V1 model (Menemenlis et al., 2008; Dutkiewicz et al., 2009; Brix et al., 2015) estimates are used as the prior ocean CO_2 exchange for all inversions, and prior errors were taken to be 100% of the flux. Fossil fuel, biofuel, and biomass burning CO_2 emissions are prescribed using the Open-source Data Inventory for Anthropogenic CO_2, version 2018 (Oda & Maksyutov, 2011; Oda et al., 2018) with downscaling to hourly emissions based on Nassar et al. (2013), CASA-GFED4-FUEL, and Global Fire Emission Database, version 4 (GFED4) (Randerson et al., 2018) inventories, respectively.

Prior error covariance matrices are taken to be diagonal, such that there are no spatial or temporal covariances. The prior NEE errors are generated based on the NEE fluxes provided by the models. It is first taken to be 60% of the NEE flux. This is then increased by scaling up the errors at times and grid cells that have active vegetation but small net fluxes. For example, the uncertainty is scaled up during the spring (source to sink) and fall (sink to source) transition periods when the 14-day NEE flux is small but the summer 14-day NEE fluxes are much larger. We also inflate the uncertainty for gridcells in which the flux is small for a given model but is much larger for the other models. The final errors range from 100% to 500% of the NEE flux. Additional details are provided in Appendix 7.

A series of flux inversions are performed that assimilate different datasets. This allows us to quantify the influence of different observational datasets on the posterior fluxes. We perform flux inversions that assimilate only ground-based in situ and flask measurements (referred to as surface-only), only TCCON measurements (TCCON-only), only GOSAT data (referred to as GOSAT-only), and all datasets simultaneously (referred to as GOSAT+surface+TCCON). For each data assimilation set-up, we perform flux inversions with each of the three prior NEE fluxes and errors. Therefore, we perform a total of 12 flux inversions.

4 Results

4.1 Evaluation of posterior-NEE-simulated CO_2

Large spatial structures in the posterior-simulated-CO_2 fields are compared with GOSAT and OCO-2 X_{CO_2}, while the accuracy of the fluxes are evaluated against aircraft-based CO_2 measurements. Rather than describing the data–model differences for all 12 inversions, the posterior fluxes are grouped by the dataset assimilated and the mean posterior fluxes are evaluated. Tables giving the data–model mismatch between the individual flux inversions and aircraft measurements are provided as supplementary materials (Tables SS1 and SS2).

4.1.1 Comparison of posterior CO_2 *against space-based* X_{CO_2}

Space-based X_{CO_2} measurements have broad spatial coverage on the timescale of a month. This allows for comparisons between modeled and measured X_{CO_2} data over large spatial scales. Here, the data–model mismatch between the posterior CO_2 fields and space-based measurements from GOSAT and OCO-2 are examined. Figure 1 shows the zonal mean data–model mismatch as a function of latitude and time for the mean prior fluxes and mean posterior fluxes for the TCCON-only inversions, surface-only inversions, GOSAT-only inversions, and GOSAT+surface+TCCON inversions. Note that there are gaps due to GOSAT's observational coverage in the tropics and at high latitudes. The mean prior flux gives larger data–model standard deviations against GOSAT (0.59 ppm) and OCO-2 (0.67 ppm) than all of the flux inversions, implying that the flux inversions improve the variance of the data–model mismatch. The CO_2 fields simulated with the prior fluxes tend to be biased low relative to GOSAT and OCO-2 during the winter and spring and biased high during the summer and fall in the northern extratropics, suggesting that the prior fluxes underestimate the magnitude of the seasonal cycle. Comparing the posterior CO_2 fields against GOSAT, the surface-only and TCCON-only flux inversions give the largest mean data–model standard deviations, which is expected as there were the only inversions that do not assimilate GOSAT data.

Comparing to OCO-2, all of the flux inversions give similar differences. Mean differences range from -0.11 ppm to 0.07 ppm and standard deviations range over 0.41-0.48 ppm, suggesting that all of the flux inversions recover the global X_{CO_2} fields with similar accuracy and precision. However, north of 40 °N, the GOSAT+surface+TCCON flux inversion shows better agreement with OCO-2 (RMS=0.30 ppm) than the other flux inversions (RMS=0.36–0.41 ppm). Differences between posterior-simulated X_{CO_2} and the OCO-2 measurements are largest in the northern subtropics, where the assimilated datasets have sparse observational coverage. Thus, it is unclear whether the differences in the subtropics are due to gaps in the observational coverage or biases in the OCO-2 retrievals.

The spread in simulated X_{CO_2} among the inversions gives a metric of the precision to which the flux inversion recovers atmospheric CO_2. Figure 2 shows the range of simulated GOSAT X_{CO_2} for the prior and posterior fluxes due to the different prior NEE fluxes and errors applied in the inversions. The largest range is obtained for the prior fluxes (mean of 1.37 ppm). The range for the TCCON-only and surface-only fluxes are reduced by 42% (0.79 ppm) and 64% (0.50 ppm) relative to the prior, respectively. However, for both flux inversions, most of the decrease in range occurs in the northern extratropics, where surface-based in situ, flask, and TCCON measurements are most concentrated. In contrast, the range increases in the tropics, where there is sparse observational coverage. This suggests that the tropical posterior NEE fluxes for the TCCON-only and surface-only flux inversions are highly sensitive to the prior NEE and error constraints. Globally, the range for GOSAT-only and GOSAT+surface+TCCON inversions are reduced by 72% and 78%, respectively, relative to the prior. The decrease relative to the prior is largest in the northern extratropics. Differences in range between the GOSAT-only and GOSAT+surface+TCCON inversions are generally quite small. The most notable differences is that the GOSAT+surface+TCCON inversions have a smaller range in the northern extratropics during the fall. GOSAT measurements do not have high sensitivity to northern extratropical fluxes during this time of year (Byrne et al., 2017), thus it appears that the surface-based measurements provide the additional information necessary to better constrain fall NEE in the northern extratropics.

4.1.2 Evaluation of posterior CO_2 *against aircraft-based measurements*

Aircraft-based measurements of atmospheric CO_2 provide a constraint on atmospheric CO_2 that is independent of the surface-based and space-based datasets assimilated. Therefore, aircraft-based CO_2 measurements offer a dataset that modeled atmospheric CO_2 can be evaluated against. Here, we evaluate the atmospheric CO_2 fields simulated using the prior and posterior fluxes against aircraft measurements over three regions with intensive sampling: East Asia, North America, and Alaska/Arctic. We only use aircraft data between 3–8 km in altitude above sea level. Differences between measured and modeled CO_2 are due to both model transport errors and surface flux errors. We have found that the differences are strongly influenced by model transport errors for individual measurements but that the impact of representativeness errors on data–model mismatches is reduced with temporal aggregation, thus we aggregate data–model mismatches to monthly means.

The GOSAT+surface+TCCON flux inversions generally show the best agreement with the aircraft-based CO_2 measurements. Figure 3 shows the monthly-mean aircraft measurements and modeled CO_2 for the three regions examined here. The GOSAT+surface+TCCON flux inversions give the smallest RMS difference against aircraft-based CO_2 in East Asia

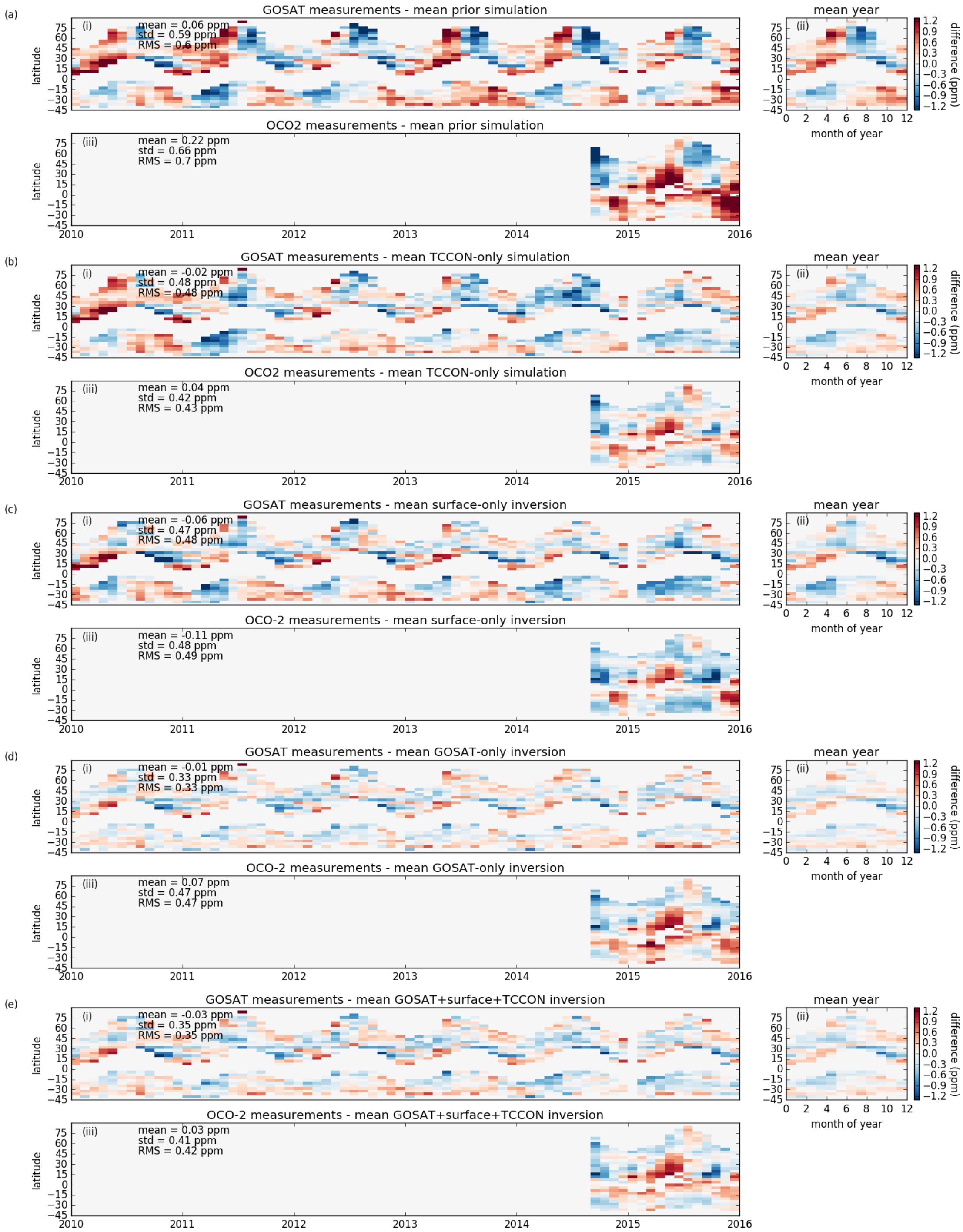

Figure 1. Zonal mean data–model mismatch for space-based X_{CO_2} measurements as a function of latitude and time for the (a) prior fluxes, (b) TCCON-only inversions (c) surface-only inversions, (d) GOSAT-only inversions, and (e) GOSAT+surface+TCCON inversions. For each set of flux inversions, the three panels show (i) the zonal and monthly mean GOSAT X_{CO_2} data–model difference for 2010 through 2015. (ii) The mean GOSAT X_{CO_2} data–model difference for each month of the year. (iii) The zonal and monthly mean OCO-2 X_{CO_2} data–model difference for 2014 through 2015.

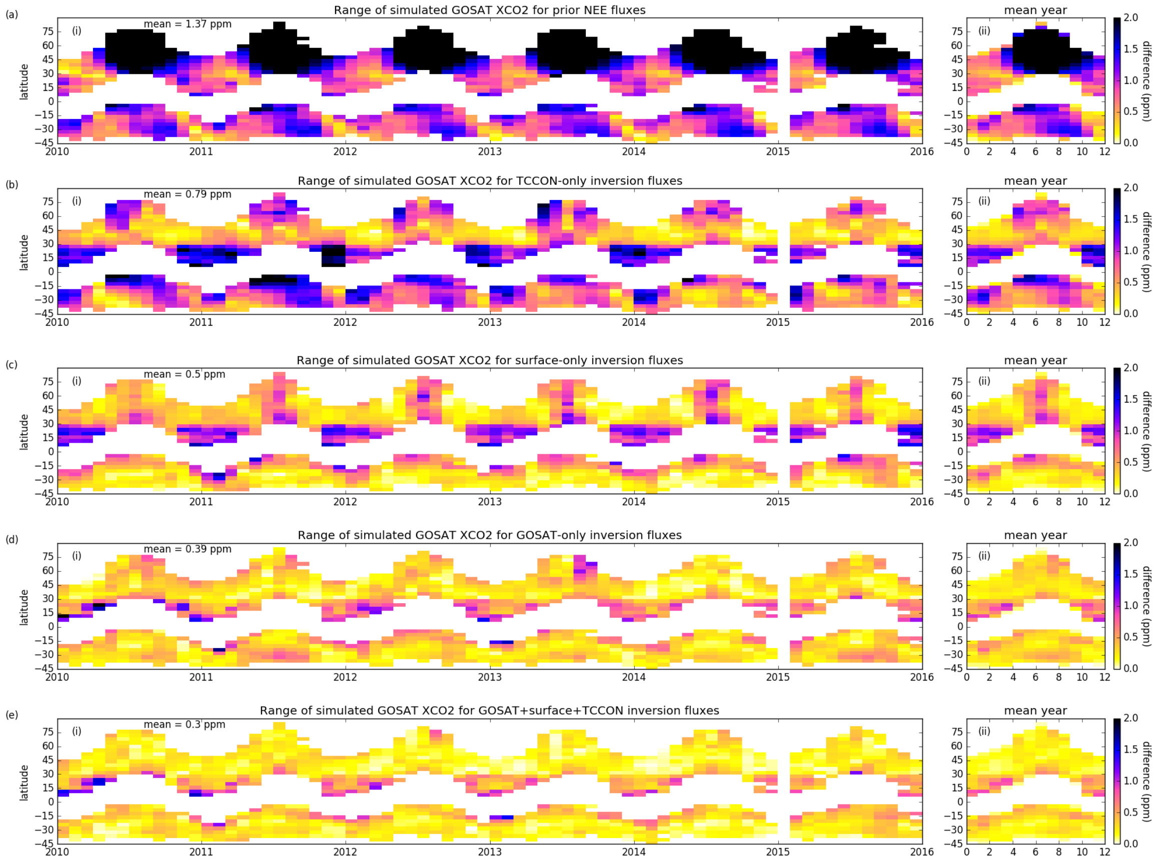

Figure 2. Spread in zonal and monthly mean simulated GOSAT X_{CO_2} for (a) prior NEE, (b) TCCON-only, (c) surface-only posterior NEE, (d) GOSAT-only posterior NEE, (e) GOSAT+surface+TCCON posterior NEE as a function of latitude and time. For each set of flux inversions sets, the panels show (i) the zonal and monthly mean range for 2010 through 2015, and (ii) The mean range for each month of the year.

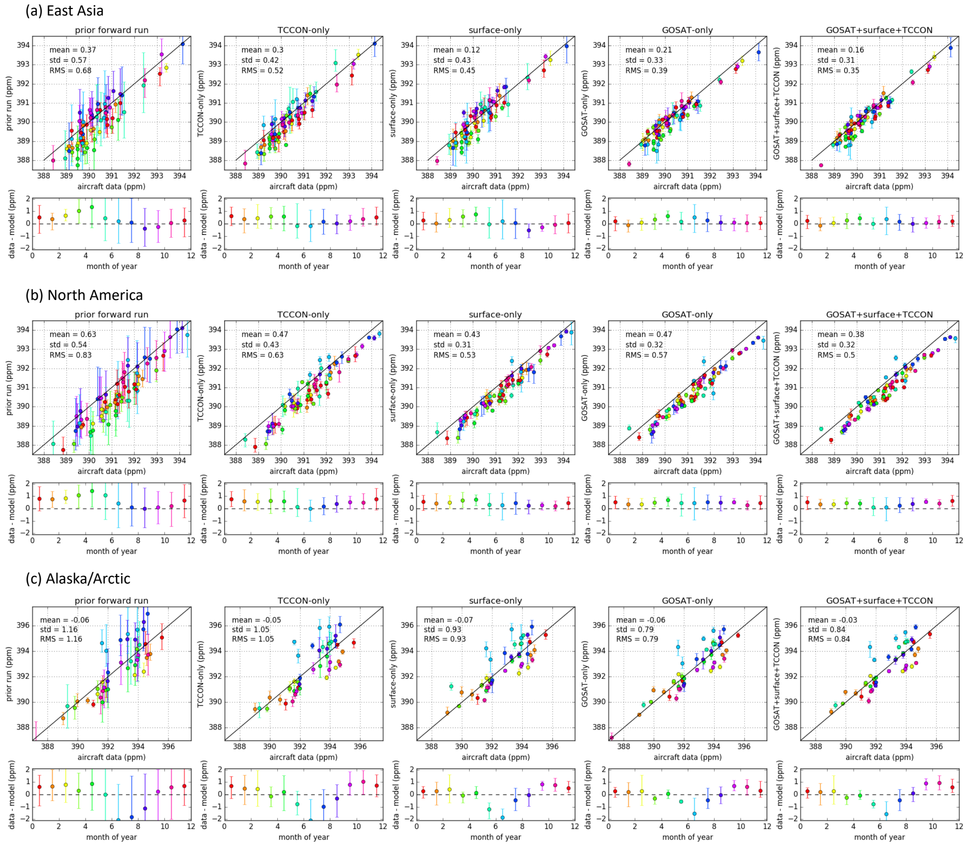

Figure 3. Comparison of monthly mean measured and simulated aircraft-based CO_2 for (a) East Asia, (b) North America, and (c) Alaska/Arctic. For each region, the mismatch for (left to right) prior, TCCON-only, surface-only, GOSAT-only, and GOSAT+surface+TCCON simulated CO_2 are shown. The top panel shows a scatter plot of the simulated aircraft-based CO_2 against the measured aircraft-based CO_2, and the error bars indicate the spread in posterior NEE. The lower panel shows the mean data–model mismatch for each month, with error bars showing the range of monthly mean mismatched over the six-years and inversion set-ups. Colors correspond to the month of year.

(0.35 ppm) and North America (0.50 ppm). The GOSAT-only flux inversions give the smallest RMS difference over the Alaska/Arctic region (0.79 ppm), although all of the flux inversions give larger RMS differences over this region relative to the midlatitude regions, suggesting that none of the flux inversions fully recover NEE at high latitudes. These aircraft measurements are also sensitive to fluxes over Siberia (Fig. S4), which is poorly observed by all datasets. Differences in the data–model mismatch between flux inversions are evident as a function of month-of-year. The GOSAT+surface+TCCON flux inversion tends to best capture month-to-month variability, while both flux inversions assimilating GOSAT measurements tend to have less seasonality in the data–model mismatch than the TCCON-only and surface-only flux inversions. This is most evident for East Asia and suggests that the GOSAT-only flux inversions better capture the month-to-month variability in fluxes (consistent with the results of Polavarapu et al. (2018) and Byrne et al. (2019)).

Despite these differences, the data–model biases against the aircraft-based measurements are generally similar between flux inversions. For example, all of the flux in-

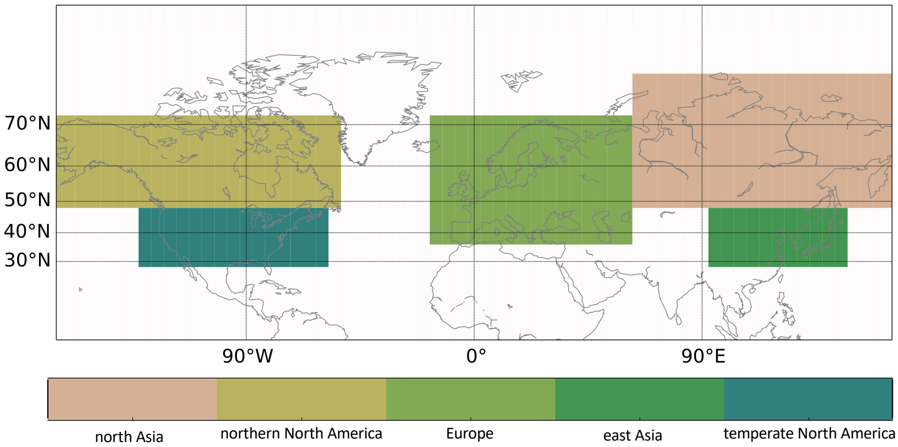

Figure 4. Five regions examined in this study. From left to right, the highlighted regions are referred to as northern North America, temperate North America, Europe, north Asia, and east Asia.

versions give positive biases for East Asia (0.12–0.30 ppm) and North America (0.38–0.47 ppm) but negative biases for the Alaska/Arctic region (-0.07 to -0.03 ppm). The fact that the data–model biases are similar suggests that these biases are sensitive to transport errors. This was quantified by regridding the fluxes and performing the evaluation against aircraft measurements at $2°\times2.5°$ spatial resolution (Figure S3). We find that model-data biases for the flux inversions change by 0.01–0.03 ppm for East Asia, 0.07–0.10 ppm for North America, and 0.08–0.11 ppm for Alaska/Arctic. These differences are similar to the magnitude of data–model differences between flux inversions, suggesting that transport model errors limit the ability of evaluating CO_2 flux estimates with aircraft-based measurements.

4.2 Mean fluxes

4.2.1 Seasonal Cycle

In the northern extratropics, the seasonal cycle of NEE produces a large annual oscillation in atmospheric CO_2, giving seasonal variations of $\sim$10 ppm in X_{CO_2}. This provides the largest signal of ecosystem carbon dynamics in atmospheric CO_2 and is the NEE signal that is best captured in CO_2 flux inversions. In this section, we examine the seasonal cycle of NEE recovered by the flux inversions in the northern extratropics grouped by the assimilated dataset. Figure 5 shows the seasonal cycle for the entire northern extratropics and five sub-continental regions (the spatial extent of the sub-continental regions are shown in Fig. 4). We examine (1) the consistency in the seasonal cycle between the datasets and (2) the precision of the posterior fluxes due to prior assumptions.

The posterior seasonal cycles of the flux inversions show consistent seasonal cycles for all assimilated datasets, relative to the prior fluxes. The GOSAT+surface+TCCON NEE fluxes most closely match the GOSAT-only NEE fluxes during the summer, as GOSAT has dense observational coverage. During the winter, the GOSAT+surface+TCCON NEE fluxes most closely match the surface-only fluxes, particularly over temperate North America and Europe where the surface-based measurements are most densely concentrated.

The spread for each set of flux inversions shows the range in posterior fluxes due to differences in the prior fluxes and errors applied. This provides a metric of the pre-

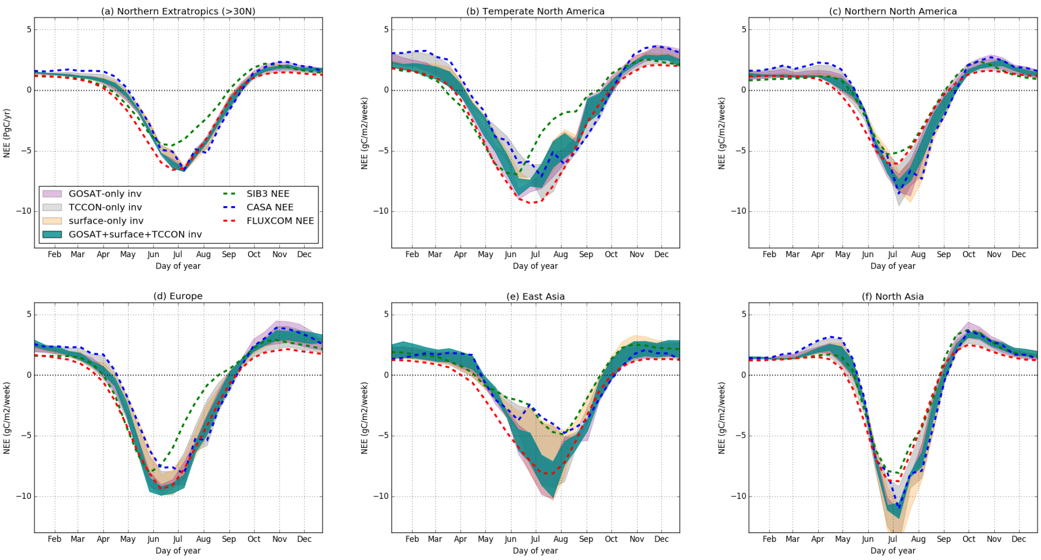

Figure 5. Prior and posterior NEE fluxes for (a) the entire northern extratropics (¿30° N), (b) temperate North America, (c) northern North America, (d) Europe, (e) east Asia, and (f) north Asia at 14 day temporal resolution. The shaded curves show the range of posterior fluxes obtained by the GOSAT-only (purple), TCCON-only (grey), surface-only (yellow), and GOSAT+surface+TCCON (dark green) flux inversions. Dashed lines show the seasonal cycles for the three prior NEE fluxes used in inversions: SiB3 (green), CASA (blue), and FLUXCOM (red).

cision to which the assimilated observations can constrain NEE. The spread is generally largest for the surface-only flux inversions outside of the winter. This is particularly notable over East Asia, where there is comparatively sparse observational coverage leading to a large spread among surface-only flux inversions. The spread is smallest for the GOSAT+surface+TCCON flux inversion, as expected. The small spread for the GOSAT+surface+TCCON flux inversions shows that the observational constraints provided by combining GOSAT, TCCON, and surface in situ and flask CO_2 measurements are sufficient to constrain the seasonal cycle of NEE on these sub-continental scales. These results suggest that the seasonal cycle is recovered by top-down flux inversions and suggests that analysis of the seasonal cycle of NEE, such as that presented by Byrne et al. (2018), could be extended to these regional scales.

4.2.2 Annual net fluxes

Here, we examine the annual net fluxes obtained for the flux inversions over the northern extratropics. Figure 6 shows the six-year mean annual net fluxes for each sub-continental region. Over the entire northern extratropics (>30° N), the flux inversions show high consistency relative to the spread in the prior. We obtain a mean annual net flux of -2.80 PgC yr^{-1} (range of -3.43 to -2.41 PgC yr^{-1}) for the TCCON-only flux inversions, -2.76 PgC yr^{-1} (range of -3.20 to -2.49 PgC yr^{-1}) for the surface-only flux inversions, -2.89 PgC yr^{-1} (range of -3.31 to -2.65 PgC yr^{-1}) for the GOSAT-only flux inversions, and -3.02 PgC yr^{-1} (range of -3.21 to -2.89 PgC yr^{-1}) for the GOSAT+surface+TCCON flux inversions. It is notable that the prior assumptions applied to the flux inversions introduce substantial differences into the posterior fluxes. The range in the northern extratropical sink due to applying different prior NEE fluxes and errors is 0.32–1.03 PgC yr^{-1}, depending on the assimilated dataset.

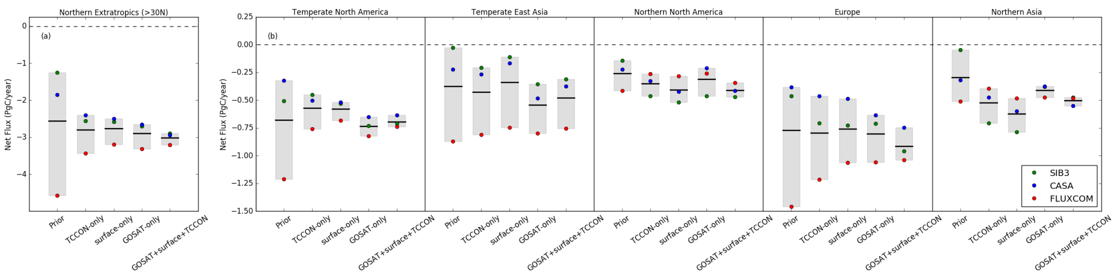

Figure 6. Six-year-mean annual net NEE fluxes for (a) all of the northern extratropics and (b) the five regions examined in this study. Shaded grey regions show the range for the prior and posterior fluxes, while the solid black line shows the mean. Individual inversions are shown by the filled circles, with colors indicating prior NEE applied: green circles indicate SiB3, blue circles indicate CASA, and red circles indicate FLUXCOM.

On regional scales, there is generally overlap in the range of net annual fluxes between the TCCON-only, surface-only, GOSAT-only, and GOSAT+surface+TCCON flux inversions. This suggests that these observational datasets provide a consistent constraint on regional net annual NEE, within the considerable uncertainty introduced through prior assumptions. The exception is north Asia, where the surface-only inversions suggest a systematically larger sink than the GOSAT-only flux inversions. This region has poor observational coverage, which may explain the differences seen here.

4.3 Interannual variability

Interannual variability (IAV) in NEE provides a measure of the response of ecosystems to climate variability. Here, we examine the IAV recovered by the flux inversions, where IAV is calculated to be the anomaly from the six-year mean. Figure 7 shows the IAV in NEE for the entire northern extratropics and five extratropical regions at 14-day temporal resolution, after performing a 3-point (42-day) running mean to filter out high frequency variability. The posterior NEE IAV is not sensitive to the prior NEE constraints applied in the flux inversion, such that similar posterior NEE IAV is recovered for each set of prior fluxes when a given assimilated dataset. This is illustrated by the small range obtained for each set of colored curves. However, the posterior NEE IAV is sensitive to the assimilated dataset, such that we find disagreement in NEE IAV for the TCCON-only, surface-only, and GOSAT-only flux inversions.

Differences in IAV between flux inversions can partially be explained by differences in the observational coverage of the datasets. As an example, let's consider the differences in IAV between the surface-only and GOSAT-only flux inversions in 2011 over temperate North America (Fig. 8). Figure 8a shows the monthly CO_2 anomalies observed by GOSAT and the surface in situ and flask network over the summer of 2011. GOSAT X_{CO_2} measurements are distributed uniformly across North America, while surface in situ and flask measurements are located south of Lake Superior. This observational coverage is reflected in the posterior fluxes. The GOSAT-only posterior NEE anomalies (Fig. 8b) reflect the large scale structures in the X_{CO_2} anomalies but miss smaller scale structures, such as the positive anomalies over south central North America. The surface-only posterior anomalies (Fig. 8c) capture large anomalies seen in CO_2, such as the anomolous release of CO_2 in south central North America, but miss much of the large scale structures. Combining these two datasets in a single inversion, referred to as "GOSAT+surface", captures both the large scale structures from the GOSAT-only and small-scale structures from the surface-only flux inversion (Fig. 8d). The posterior NEE anomalies from the

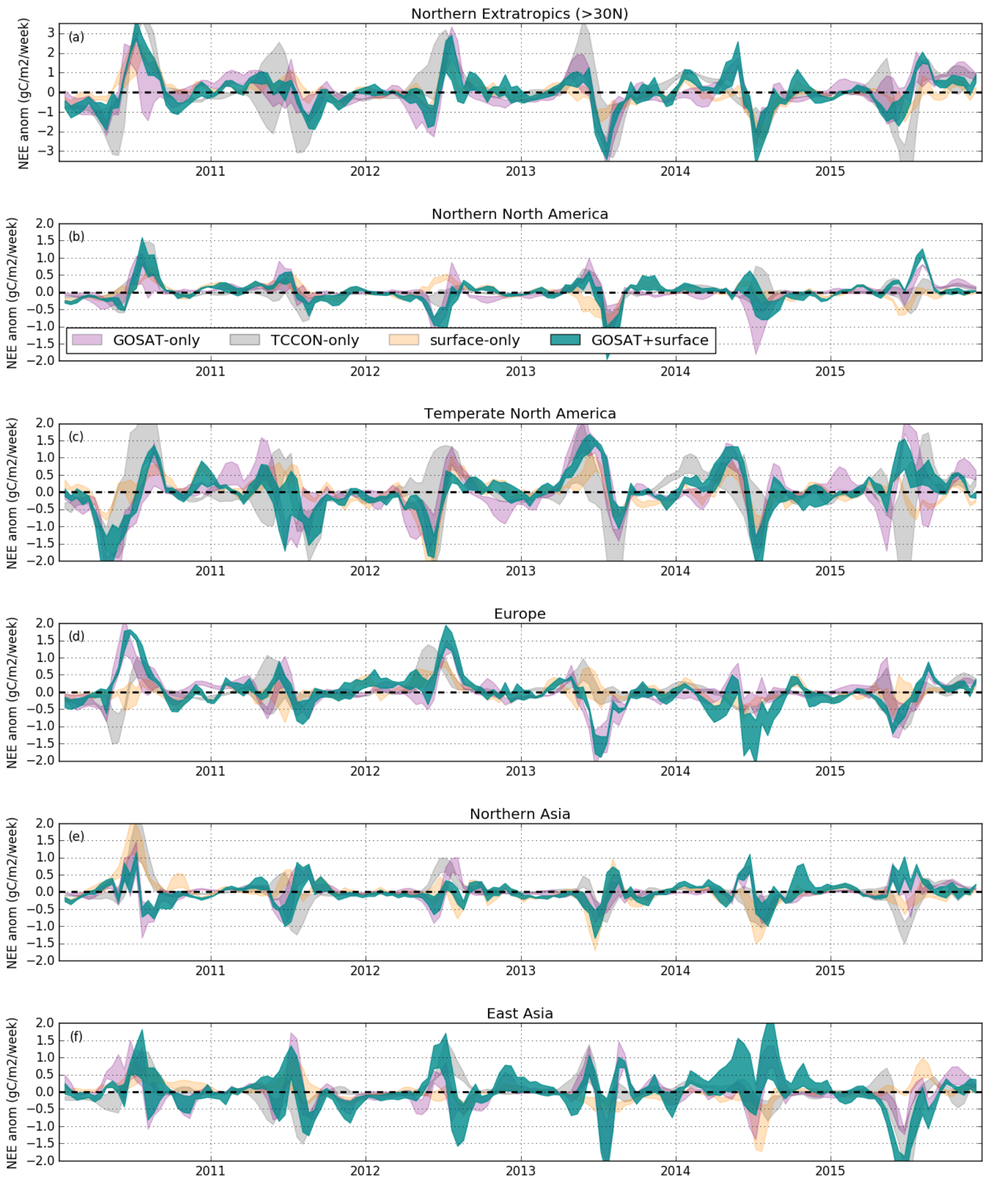

Figure 7. IAV in NEE for 2010–2015 at 14 day temporal resolution for (a) the entire northern extratropics (¿30° N), (b) temperate North America, (c) northern North America, (d) Europe, (e) east Asia, and (f) north Asia. The shaded curves show the range of posterior fluxes obtained by the GOSAT-only (purple), TCCON-only (grey), surface-only (yellow), and GOSAT+surface+TCCON (dark green) flux inversions. A 3-point (42 day) running mean is performed to remove high frequency variability.

GOSAT+surface flux inversion also correlate with anomalies in soil temperature (mean of MERRA-2 soil temperature over levels 1–3, Reichle et al. (2011, 2017)) (Fig. 8e) and soil moisture (ESA CCI Surface Soil Moisture Product, Y. Y. Liu et al. (2011, 2012); Wagner et al. (2012); Gruber et al. (2017); Dorigo et al. (2017)) (Fig. 8f) over this time period, suggesting that combining these datasets produces more realistic NEE IAV. Similar results were found over Eurasia during the summer of 2010 (Fig. S5).

On an annual basis, we find mixed agreement between flux inversions in year-to-year variations. Figure 9 shows IAV in annual net NEE anomalies for the entire northern extratropics. In general, IAV in annual net fluxes are consistent for a given set of assimilated data, suggesting that the results are not sensitive to the prior fluxes and errors used. Note that the prior NEE fluxes did not contain IAV, which has previously been shown to have a substantial impact on posterior NEE IAV (Byrne et al., 2019). However, posterior IAV is quite variable between different assimilated datasets. The cause of these differences between the flux inversions are likely partially due to differences in the observational coverage between datasets. It is possible that differences between datasets are also partially due to changes in the observational coverage over time, which has previously been shown to have an impact on inferred fluxes (Rödenbeck et al., 2003; Gurney et al., 2008; Bruhwiler et al., 2011).

5 Discussion

5.1 Consistency in surface-based and spaced-based flux constraints

The results generally show good agreement between the flux inversions assimilating different datasets. The agreement between the surface-only and GOSAT-only flux inversions may seem surprising in the context of a number of previous studies that have shown substantial differences between surface-based and space-based flux estimates (Basu et al., 2013; Chevallier et al., 2014; Houweling et al., 2015). However, more recent studies have shown improved agreement between surface-based and space-based flux inversions. Chevallier et al. (2019) found that flux inversions assimilating OCO-2 ACOS version 9 measurements gave similar net annual fluxes to those assimilating surface-based measurements, and that both compared well against aircraft measurements. Interestingly, Chevallier et al. (2019) also found that GOSAT OCO Full Physics (OCFP) v7.1 XCO_2 retrievals did not compare as well against aircraft measurements. Comparisons between the ACOS 7.3 and OCFP v7.1 (downloaded from the Copernicus Climate Change Service, https://climate.copernicus.eu/) show substantial differences in zonal mean X_{CO_2} (Fig. S6). Furthermore, GOSAT ACOS 7.3 retrievals are found to give better agreement with posterior-simulated-CO_2 from the surface-only flux inversion (Fig. S7). This suggests that the specific retrieval algorithm used has a large impact on the posterior fluxes, such that the improved agreement between surface-based and space-based measurements found in recent studies may be primarily due to improvements in the ACOS X_{CO_2} retrieval algorithm. Miller and Michalak (2019) have also argued that recent improvements in the ACOS algorithm have substantially increased the reliability of OCO-2 X_{CO_2} measurements in flux inversions studies (for version 8 in particular). Substantial work has gone into refining the ACOS retrieval algorithm over the past decade (O'Dell et al., 2012; Crisp et al., 2012; Eldering et al., 2017; O'Dell et al., 2018; Kiel et al., 2019; Nelson & O'Dell, 2019). Thus, the improved agreement between surface-based and space-based CO_2 constraints is likely best explained by improvements in the ACOS retrieval algorithm.

A consistent six-year mean northern extratropical sink is obtained by all observational datasets. This result is in contrast to several previous studies that found substantial differences in the annual net NEE flux of CO_2 in the northern extratropics between flux inversions assimilating surface-based and space-based measurements (Basu et al., 2013; Saeki et al., 2013; Chevallier et al., 2014; Reuter et al., 2014). The reason why we obtain a more consistent annual net flux between datasets than some earlier studies is

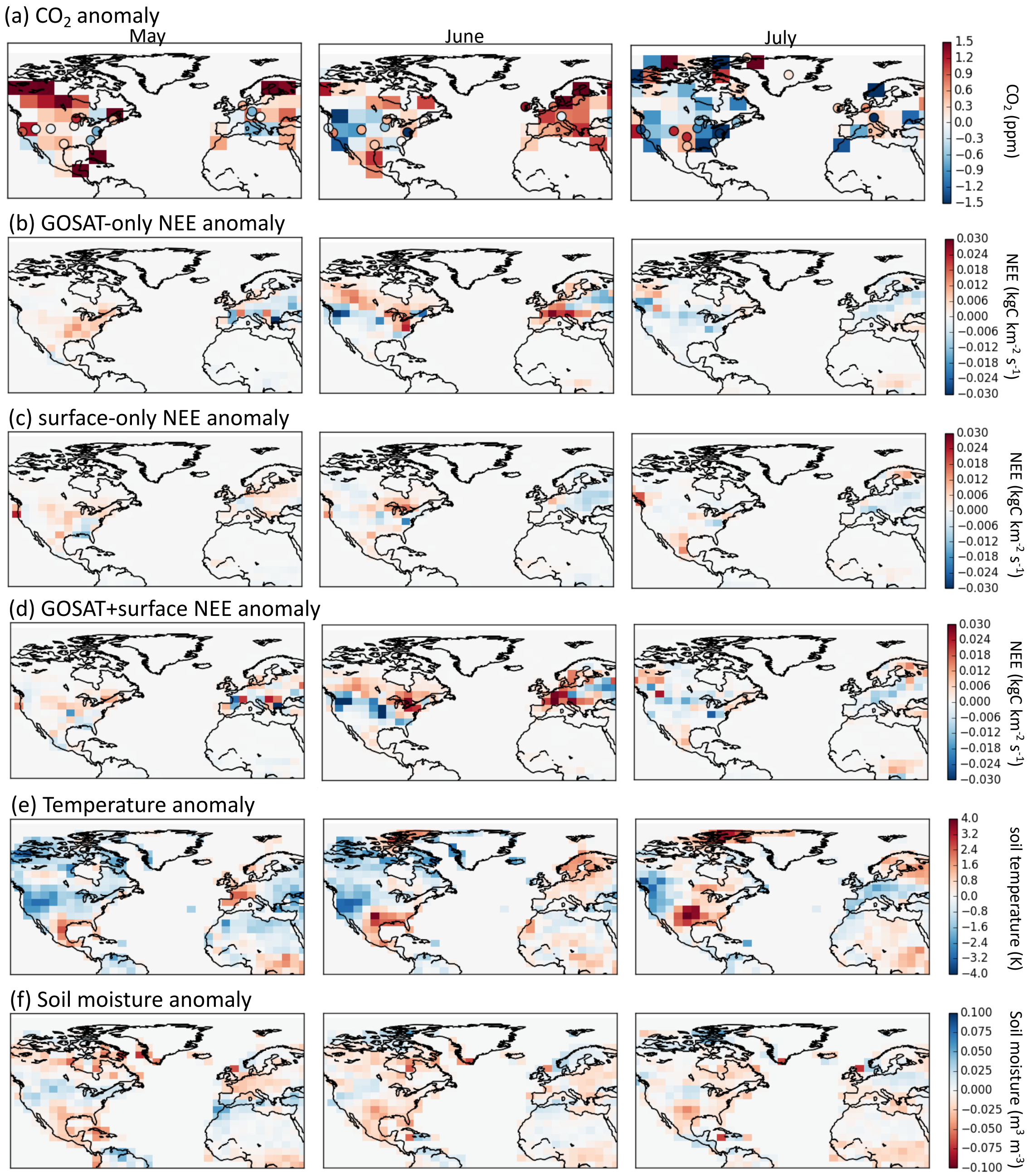

Figure 8. Monthly anomalies in (a) GOSAT X_{CO_2} (ppm, 4° × 5° grid cells) and surface site CO_2 (ppm divided by four, circles), (b) GOSAT-only posterior NEE, (c) surface-only posterior NEE, (d) GOSAT+surface posterior NEE, (e) MERRA-2 soil temperature, (f) ESA CCI soil moisture, for (left-to-right) May, June, and July of 2011.

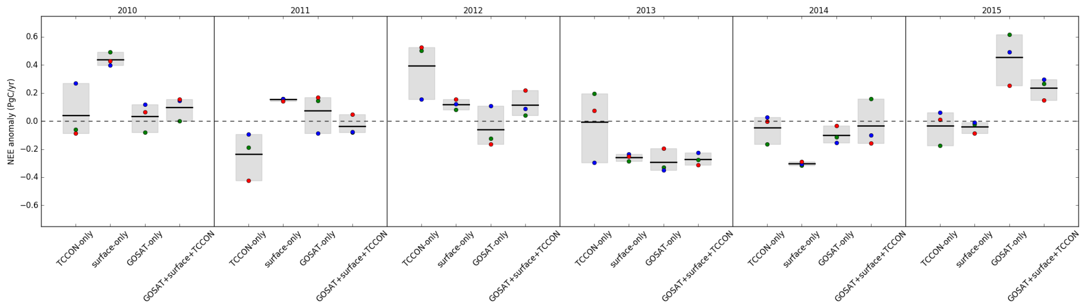

Figure 9. Annual net IAV in NEE over 2010-2015 for the TCCON-only, surface-only, GOSAT-only, and GOSAT+surface+TCCON flux inversions. Shaded grey regions show the range for the fluxes, while the solid black line shows the mean. Individual inversions are shown by the filled circles, with colors indicating prior NEE applied: green circles indicate SiB3, blue circles indicate CASA, and red circles indicate FLUXCOM.

not immediately clear, but could be due to advancements in the retrieval algorithm (e.g., ACOS 3.3 and earlier versions were used in Houweling et al. (2015)) or due to the fact that we look at a multi-year mean while earlier studies looked at shorter time periods (e.g., Houweling et al. (2015) only examined June 2009 to June 2010). In fact, we find that the surface-only inversion suggests weaker uptake in 2010 than average (by 0.40 to 0.49 PgC yr^{-1}), while the GOSAT flux inversion suggests near average uptake (see Sec. 4.3), suggesting that the difference in inferred fluxes between these two datasets may have been unusually large for 2010. However, it is important to note that differences in annual net fluxes do not imply biases in the measurements. There are aspects of the inversion set-ups that can lead to differences. For example, differences in the distribution of observations can lead to significant differences in annual net fluxes (J. Liu et al., 2014; Byrne et al., 2017; Basu et al., 2018). Thus, one should not necessarily expect consistent annual net fluxes from observational datasets with spatial and temporal gaps in observational coverage.

5.2 Does combining datasets improve flux inversions?

Is it possible to conclude that the GOSAT+surface+TCCON flux inversions improve flux estimates relative to the flux inversions that assimilate a single dataset? Of course, the answer to this question depends on how “improve” is defined. The GOSAT+surface+TCCON flux inversions generally show a small reduction in model-data differences against independent aircraft-based CO_2 and OCO-2 X_{CO_2} (north of 40°N). This suggests that combining these datasets in a flux inversion framework produces NEE fluxes that better recover the true atmospheric CO_2 fields than any dataset alone. However, confounding factors in evaluating these fluxes remain a significant concern. Model transport errors appear to be a main driver of data-model differences for aircraft-based CO_2 measurements, and obscures the source of data–model differences. Evaluating optimized fluxes against OCO-2 is also problematic because these retrievals are known to have their own biases.

The GOSAT+surface+TCCON flux inversions improve the precision of the posterior NEE fluxes relative to the flux inversions assimilating one dataset. This is found to be the case at seasonal, annual, and interannual scales. The GOSAT+surface+TCCON flux inversions closely resemble the GOSAT-only NEE fluxes during the summer and surface-only fluxes during the winter for five northern extratropical regions. This is expected given the spatiotemporal distribution of GOSAT and surface-based CO_2 measurements and suggests that the GOSAT+surface+TCCON posterior NEE fluxes are better constrained by the observations than the GOSAT-only or surface-only flux inversions. Therefore, the GOSAT+surface+TCCON flux inversions are less likely to be impacted by biases in the observational coverage, such that, from an observational coverage perspective, we can conclude that the GOSAT+surface+TCCON flux inversions are better constrained than the GOSAT-only or surface-only flux inversions.

An important concern in combining CO_2 datasets within a single flux inversion system is that there could be relative biases in the atmospheric CO_2 constraints provided by the different datasets. Any inconsistency in flux constraints between datasets has the potential of introducing artifacts into the posterior fluxes. Biases in the observations could be present due to errors in the X_{CO_2} retrieval algorithm, representativness errors (Agustí-Panareda et al., 2019) or model transport errors. Several previous studies have suggested that unrealistically large uptake over Europe ($\sim$1.5 PgC yr^{-1}) is recovered in posterior fluxes due to biases in the GOSAT retrieval algorithm (Basu et al., 2013; Chevallier et al., 2014), although the ACOS retrieval algorithm has undergone significant development since these studies (Eldering et al., 2017; O'Dell et al., 2018) resulting in reduced biases (Miller & Michalak, 2019). Similarly, a number of studies have pointed out systematic transport errors in GEOS-Chem (Yu et al., 2018; Schuh et al., 2019), as-well as biases in reanalysis winds (e.g., vertical mixing, Parazoo et al. (2012)). We do not find clear evidence for biases between the surface-based and GOSAT constraints, although, these biases may be challenging to identify. However, we do see the impact of model transport errors in comparisons between the posterior-simulated-CO_2 and aircraft measurements. Ideally, this analysis should be performed with two different transport models so that transport related errors could be more easily identified.

6 Conclusions

This study presented a series of flux inversions assimilating surface-based flask and in situ CO_2 measurements, TCCON X_{CO_2}, GOSAT X_{CO_2}, or all datasets combined. All of the flux inversions showed improved agreement with independent aircraft-based CO_2 measurements relative to prior flux estimates. The GOSAT+surface+TCCON flux inversion gave the smallest RMS differences against aircraft-based CO_2 measurements over East Asia and North America, and OCO-2 X_{CO_2} measurements (north of 40° N), suggesting that combining the datasets improves flux estimates. However, the data–model mismatches were strongly impacted by transport model, which makes robust evaluations of posterior surface fluxes challenging.

We found that all observing systems generally give consistent posterior NEE fluxes relative to the spread in prior fluxes. This suggests that these datasets provide consistent information on NEE. The GOSAT+surface+TCCON posterior NEE most closely resembles the GOSAT-only posterior NEE during the summer and surface-only posterior NEE during the winter, consistent with the temporal variations in the observational constraints. This suggests that the GOSAT+surface+TCCON flux inversions benefit from the improved spatiotemporal distribution of measurements, providing posterior fluxes that are better informed by measurements throughout the year.

The results of this study suggest that surface-based and space-based atmospheric CO_2 constraints provide consistent constraints on NEE fluxes, and can be combined in a flux inversion framework. This result stands in contrast to earlier attempts to combine these datasets (Houweling et al., 2015), and suggests that the improved consistency between the datasets has been made possible by the considerable effort spent refining the ACOS retrieval algorithm (Eldering et al., 2017; O'Dell et al., 2018; Kiel et al., 2019; Chevallier et al., 2019; Miller & Michalak, 2019).

7 Appendix: Prior NEE fluxes and errors

7.1 Simple biosphere model (SiB3)

SiB3 was originally designed as a lower boundary for General Circulation Models with explicit treatment of biophysical processes. The ability to ingest satellite phenology was later introduced (P. Sellers et al., 1996; P. J. Sellers et al., 1996), and further refinements included a prognostic canopy air space (Vidale & Stöckli, 2005), more realistic soil and snow (I. Baker et al., 2003) and modifications to calculations of root water uptake and soil water stress (I. Baker et al., 2008). The current version is called SiB3. Simulations used in this analysis use phenology (Leaf Area Index, LAI; fraction of Photosynthetically Active Radiation, fPAR) from the Moderate Resolution Imaging Spectroradiometer (MODIS). MERRA reanalysis is used as model inputs, with precipitation scaled to Global Precipitation Climatology Project (GPCP: Adler et al. (2003)) following I. T. Baker et al. (2010).

These fluxes are adjusted to obtain a global net drawdown equal to 4.6 PgC yr^{-1}. To do this, the annual net flux at each grid cell and global total annual net drawdown are calculated. The annual net flux at each gridcell is then scaled so that the annual net flux is 4.6 PgC yr^{-1}. The difference between the original and scaled annual net flux at each grid cell is then calculated. From this difference, an adjustment at each grid cell for each 14-day period is performed so that the annual net flux then equals the scaled annual net flux at each grid cell.

The prior NEE errors are generated based on the NEE fluxes provided by the models. It is first taken to be 60% of the NEE flux. This is then increased by scaling up the errors if the mean flux for a given gridcell is large but the flux is small at a given time. For example, the uncertainty is scaled up during the fall. We also inflate the uncertainty where the flux is small for SiB3 but large for CASA and FLUXCOM. The final errors range from 100% to 500% of the NEE flux.

7.2 CASA

The version of the model used here, CASA-GFED3, was modified from Potter et al. (1993) as described in Randerson et al. (1996) and van der Werf et al. (2006). It is driven by MERRA reanalysis and satellite-observed NDVI to track plant phenology. We use the same fluxes as are used for the CarbonTracker 2016 (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/) prior. CASA outputs monthly fluxes of Net Primary Productivity (NPP) and heterotropic respiration (R_H). From these fluxes, GPP and ecosystem respiration (R_e) are estimated to be GPP = 2NPP and $R_e = R_H - \text{NPP}$. Temporal downscaling and smoothing was performed from monthly CASA fluxes to 90-min fluxes using temperature and shortwave radiation from the ECMWF ERA-interim reanalysis (note this method differs from Olsen and Randerson (2004). GFED_CMS is used for global fire emissions (http://nacp-files.nacarbon.org/nacp-kawa-01/). We use average model fluxes by averaging the fluxes for 2007–2012.

These fluxes are adjusted to obtain a global net drawdown equal to 4.6 PgC yr^{-1}. To do this, the annual net flux at each grid cell and global total annual net drawdown are calculated. The annual net flux at each gridcell is then scaled so that the annual net flux is 4.6 PgC yr^{-1}. The difference between the original and scaled annual net flux at each grid cell is then calculated. From this difference, an adjustment at each grid cell for each 14-day period is performed so that the annual net flux then equals the scaled annual net flux at each grid cell.

The prior NEE errors are generated based on the NEE fluxes provided by the models. It is first taken to be 60% of the NEE flux. This is then increased by scaling up the errors if the mean flux for a given gridcell is large but the flux is small at a given time. For example, the uncertainty is scalled up during the fall. We also inflate the uncertainty where the flux is small for CASA but large for SiB3 and FLUXCOM. The final errors range from 100% to 500% of the NEE flux.

7.3 FLUXCOM

FLUXCOM products are generated using upscaling approaches based on machine learning methods that integrate FLUXNET site level observations, satellite remote sensing, and meteorological data (Tramontana et al., 2016; Jung et al., 2017). Jung et al. (2017) generate R_e products using several machine learning methods. For this study, we downloaded the products generated using random forests (RF), multivariate regression splines (MARS) and artificial neural networks (ANN) at daily resolution from the Data Portal of the Max Planck Institute for Biochemistry (https://www.bgc-jena.mpg.de). The mean seasonal cycle over 2008-2012 is calculated for each product.

These fluxes are adjusted to obtain a global net drawdown equal to 4.6 PgC yr^{-1}. For FLUXCOM, we only adjust fluxes south of 35° N because the northern extratropical NEE fluxes have been heavily informed by FLUXNET sites. For grid cells south of 35° N, the annual net flux at each grid cell and global total annual net drawdown are calculated. The annual net flux at each gridcell is then scaled so that the annual net flux is 4.6 PgC yr^{-1}. The difference between the original and scaled annual net flux at each grid cell is then calculated. From this difference, an adjustment at each grid cell for each 14-day period is performed so that the annual net flux then equals the scaled annual net flux at each grid cell.

The prior NEE errors are generated based on the NEE fluxes provided by the models. It is first taken to be 60% of the NEE flux. This is then increased by scaling up the errors if the mean flux for a given gridcell is large but the flux is small at a given time. For example, the uncertainty is scalled up during the fall. We also inflate the uncertainty where the flux is small for FLUXCOM but large for SiB3 and CASA. The final errors range from 100% to 500% of the NEE flux.

Acknowledgments

Brendan Byrne's research was supported by an appointment to the NASA Postdoctoral Program at the Jet Propulsion Laboratory, administered by Universities Space Research Association under contract with NASA. The research carried out at the Jet Propulsion Laboratory, California Institute of Technology, was under a contract with the National Aeronautics and Space Administration. Kevin Bowman's research was supported by the NASA Carbon Monitoring System (CMS) project (NNH16ZDA001N-CMS). Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Advanced Supercomputing (NAS) Division at Ames Research Center. The JR-STATION dataset is available from the Global Environmental Database, hosted by Center for Global Environmental Research (CGER), National Institute for Environmental Studies (NIES)

(http://db.cger.nies.go.jp/portal/geds/atmosphericAndOceanicMonitoring). TCCON data were obtained from the TCCON Data Archive, hosted by CaltechDATA (http://tccondata.org). FLUXCOM products were obtained from the Data Portal of the Max Planck Institute for Biochemistry [https://www.bgc-jena.mpg.de]. MERRA-2 products were downloaded from MDISC [https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/], managed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC). GOSAT OCFP v7.1 X_{CO_2} retrievals were downloaded from the Copernicus Climate Change Service website (https://cds.climate.copernicus.eu). Version 4.1 of the GLOBALVIEW plus package was downloaded from http://www.esrl.noaa.gov/gmd/ccgg/obspack/. ESA CCI soil moisture data was downloaded from https://www.esa-soilmoisture-cci.org/. Odiac emissions dataset was provided by T. Oda of Colorado State University, Fort Collins CO, USA/Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder

CO, USA. Odiac project is supported by Greenhouse Gas Observing SATellite (GOSAT) project, National Institute for Environmental Studies (NIES), Japan. We thank S. Basu for providing downscaled ODIAC emissions. We thank T. Machida, H. Matsueda, Y. Sawa, and Y. Niwa for providing CONTRAIL measurements. The TCCON site at Réunion Island is operated by the Royal Belgian Institute for Space Aeronomy with financial support in 2014, 2015, 2016, 2017, 2018 and 2019 under the EU project ICOS-Inwire and the ministerial decree for ICOS (FR/35/IC2) and local activities supported by LACy/UMR8105 – Université de La Réunion. The TCCON project for Rikubetsu site is supported in part by the GOSAT series project.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., ... others (2003). The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, *4*(6), 1147–1167.

Agustí-Panareda, A., Diamantakis, M., Massart, S., Chevallier, F., Muñoz Sabater, J., Barré, J., ... Wunch, D. (2019). Modelling CO_2 weather – why horizontal resolution matters. *Atmos. Chem. Phys.*, *19*(11), 7347–7376. Retrieved from `https://www.atmos-chem-phys.net/19/7347/2019/` doi: 10.5194/acp-19-7347-2019

Bacastow, R. (1976). Modulation of atmospheric carbon dioxide by the southern oscillation. *Nature*, *261*(5556), 116–118. doi: 10.1038/261116a0

Baker, I., Denning, A. S., Hanan, N., Prihodko, L., Uliasz, M., Vidale, P.-L., ... Bakwin, P. (2003). Simulated and observed fluxes of sensible and latent heat and CO_2 at the WLEF–TV tower using SiB2. 5. *Glob. Change Biol.*, *9*(9), 1262–1277.

Baker, I., Prihodko, L., Denning, A., Goulden, M., Miller, S., & Da Rocha, H. (2008). Seasonal drought stress in the Amazon: Reconciling models and observations. *J. Geophys. Res.-Biogeo.*, *113*(G00B01). doi: 10.1029/2007JG000644

Baker, I. T., Denning, A. S., & Stöckli, R. (2010). North American gross primary productivity: regional characterization and interannual variability. *Tellus B*, *62*(5), 533–549. doi: 10.1111/j.1600-0889.2010.00492.x

Basu, S., Baker, D. F., Chevallier, F., Patra, P. K., Liu, J., & Miller, J. B. (2018). The impact of transport model differences on co 2 surface flux estimates from oco-2 retrievals of column average co 2. *Atmospheric Chemistry & Physics*, *18*(10).

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., ... others (2013). Global CO_2 fluxes estimated from GOSAT retrievals of total column CO_2. *Atmos. Chem. Phys.*, *13*(17), 8695–8717. doi: 10.5194/acp-13-8695-2013

Bolin, B., & Keeling, C. (1963). Large-scale atmospheric mixing as deduced from the seasonal and meridional variations of carbon dioxide. *J. Geophys. Res.*, *68*(13), 3899–3920. doi: 0.1029/JZ068i013p03899

Bowman, K., Liu, J., Bloom, A., Parazoo, N., Lee, M., Jiang, Z., ... others (2017). Global and Brazilian carbon response to El Niño Modoki 2011–2010. *Earth and Space Sci.*, *4*(10), 637–660. doi: 10.1002/2016EA000204

Brix, H., Menemenlis, D., Hill, C., Dutkiewicz, S., Jahn, O., Wang, D., ... Zhang, H. (2015). Using green's functions to initialize and adjust a global, eddying ocean biogeochemistry general circulation model. *Ocean Modelling*, *95*, 1–14.

Bruhwiler, L., Michalak, A., & Tans, P. (2011). Spatial and temporal resolution of carbon flux estimates for 1983–2002. *Biogeosciences*, *8*(5), 1309–1331. doi: 10.5194/bg-8-1309-2011

Byrne, B., Jones, D. B. A., Strong, K., Polavarapu, S. M., Harper, A. B., Baker, D. F., & Maksyutov, S. (2019). On what scales can gosat flux inversions constrain anomalies in terrestrial ecosystems? *Atmos. Chem. Phys.*, *19*(20), 13017–13035. doi: 10.5194/acp-19-13017-2019

Byrne, B., Jones, D. B. A., Strong, K., Zeng, Z.-C., Deng, F., & Liu, J. (2017). Sensitivity of CO_2 surface flux constraints to observational coverage. *J. Geophys. Res.-Atmos*, *112*(12), 6672–6694. doi: 10.1002/2016JD026164

Byrne, B., Wunch, D., Jones, D., Strong, K., Deng, F., Baker, I., … others (2018). Evaluating GPP and respiration estimates over northern midlatitude ecosystems using solar-induced fluorescence and atmospheric CO_2 measurements. *Journal of Geophysical Research: Biogeosciences*, *123*(9), 2976–2997.

Chatterjee, A., Gierach, M., Sutton, A., Feely, R., Crisp, D., Eldering, A., … Schimel, D. (2017). Influence of El Niño on atmospheric CO_2 over the tropical Pacific Ocean: Findings from NASA's OCO-2 mission. *Science*, *358*(6360), eaam5776. doi: 10.1126/science.aam5776

Chevallier, F., Deutscher, N. M., Conway, T., Ciais, P., Ciattaglia, L., Dohe, S., … others (2011). Global CO_2 fluxes inferred from surface air-sample measurements and from TCCON retrievals of the CO_2 total column. *Geophys. Res. Lett.*, *38*(24).

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

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., & Cozic, A. (2019). Objective evaluation of surface- and satellite-driven co_2 atmospheric inversions. *Atmos. Chem. Phys. Discuss.*, *2019*, 1–28. doi: 10.5194/acp-2019-213

Cooperative Global Atmospheric Data Integration Project. (2018). *Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2017; obspack_CO_2_1_globalviewplus_v4.1_2018_10_29; noaa earth system research laboratory, global monitoring division.* doi: 10.25925/20181026

Crisp, D., Atlas, R., Breon, F.-M., Brown, L., Burrows, J., Ciais, P., … others (2004). The orbiting carbon observatory (OCO) mission. *Advances in Space Research*, *34*(4), 700–709.

Crisp, D., Fisher, B. M., O'Dell, C., Frankenberg, C., Basilio, R., Bösch, H., … Yung, Y. L. (2012). The ACOS CO_2 retrieval algorithm-Part II: Global XCO_2 data characterization. *Atmos. Meas. Tech.*, *5*(4), 687–707. doi: 10.5194/amt-5-687-2012

De Mazière, M., Sha, M. K., Desmet, F., Hermans, C., Scolas, F., Kumps, N., … Cammas, J.-P. (2017). *Tccon data from réunion island (re), release ggg2014.r1.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/322` doi: 10.14291/tccon.ggg2014.reunion01.r1

Diallo, M., Legras, B., Ray, E., Engel, A., & Añel, J. A. (2017). Global distribution of CO_2 in the upper troposphere and stratosphere. *Atmos. Chem. Phys.*, *17*(6), 3861–3878. Retrieved from `https://www.atmos-chem-phys.net/17/3861/2017/` doi: 10.5194/acp-17-3861-2017

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., … others (2017). Esa cci soil moisture for improved earth system understanding: State-of-the art and future directions. *Remote Sens. Environ.*, *203*, 185–215.

Dutkiewicz, S., Follows, M. J., & Bragg, J. G. (2009). Modeling the coupling of ocean ecology and biogeochemistry. *Global Biogeochem. Cy.*, *23*(4).

Eldering, A., O'Dell, C. W., Wennberg, P. O., Crisp, D., Gunson, M. R., Viatte, C., … Yoshimizu, J. (2017). The orbiting carbon observatory-2: first 18 months of science data products. *Atmos. Meas. Tech.*, *10*(2), 549–563. doi: 10.5194/amt-10-549-2017

Feist, D. G., Arnold, S. G., John, N., & Geibel, M. C. (2017). *Tccon data from ascension island (sh), release ggg2014.r0.* CaltechDATA. Retrieved from https://data.caltech.edu/records/210 doi: 10.14291/tccon.ggg2014.ascension01.r0/1149285

Fischer, M. L., Parazoo, N., Brophy, K., Cui, X., Jeong, S., Liu, J., ... others (2017). Simulating estimation of california fossil fuel and biosphere carbon dioxide exchanges combining in situ tower and satellite column observations. *J. Geophys. Res.-Atmos.*, *122*(6), 3653–3671.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., ... others (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Climate*, *30*(14), 5419–5454.

Griffith, D. W., Deutscher, N. M., Velazco, V. A., Wennberg, P. O., Yavin, Y., Keppel-Aleks, G., ... Bryant, G. W. (2017). *Tccon data from darwin (au), release ggg2014.r0.* CaltechDATA. Retrieved from https://data.caltech.edu/records/269 doi: 10.14291/tccon.ggg2014.darwin01.r0/1149290

Griffith, D. W., Velazco, V. A., Deutscher, N. M., Paton-Walsh, C., Jones, N. B., Wilson, S. R., ... Riggenbach, M. O. (2017). *Tccon data from wollongong (au), release ggg2014.r0.* CaltechDATA. Retrieved from https://data.caltech.edu/records/291 doi: 10.14291/tccon.ggg2014.wollongong01.r0/1149291

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple collocation-based merging of satellite soil moisture retrievals. *IEEE T. Geosci. Remote*, *55*(12), 6780–6792.

Gurney, K. R., Baker, D., Rayner, P., & Denning, S. (2008). Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated from atmospheric CO_2 inversions for the period 1980 to 2005. *Global Biogeochem. Cy.*, *22*(3). Retrieved from http://dx.doi.org/10.1029/2007GB003082 (GB3025) doi: 10.1029/2007GB003082

Houweling, S., Baker, D., Basu, S., Boesch, H., Butz, A., Chevallier, F., ... others (2015). An intercomparison of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements. *J. Geophys. Res.-Atmos.*, *120*(10), 5253–5266.

ICOS RI. (2019). *Icos atmospheric greenhouse gas mole fractions of CO_2, CH_4, CO, $^{14}CO_2$ and meteorological observations september 2015 - april 2019 for 19 stations (49 vertical levels), final quality controlled level 2 data (version 1.0). icos eric - carbon portal.* doi: 10.18160/CE2R-CC91

Iraci, L. T., Podolske, J. R., Hillyard, P. W., Roehl, C., Wennberg, P. O., Blavier, J.-F., ... Boyden, H. (2017). *Tccon data from edwards (us), release ggg2014.r1.* CaltechDATA. Retrieved from https://data.caltech.edu/records/270 doi: 10.14291/tccon.ggg2014.edwards01.r1/1255068

Jung, M., Reichstein, M., Schwalm, C. R., Huntingford, C., Sitch, S., Ahlström, A., ... others (2017). Compensatory water effects link yearly global land CO_2 sink changes to temperature. *Nature*, *541*(7638), 516–520.

Keeling, C. D. (1960). The concentration and isotopic abundances of carbon dioxide in the atmosphere. *Tellus*, *12*(2), 200–203.

Keeling, C. D., Chin, J., & Whorf, T. (1996). Increased activity of northern vegetation inferred from atmospheric CO_2 measurements. *Nature*, *382*(6587), 146-149.

Kiel, M., O'Dell, C. W., Fisher, B., Eldering, A., Nassar, R., MacDonald, C. G., & Wennberg, P. O. (2019). How bias correction goes wrong: measurement of xco2 affected by erroneous surface pressure estimates. *Atmos. Meas. Tech.*, *12*(4).

Liu, J., Bowman, K. W., Lee, M., Henze, D. K., Bousserez, N., Brix, H., ... others (2014). Carbon monitoring system flux estimation and attribution: impact of ACOS-GOSAT XCO_2 sampling on the inference of terrestrial biospheric sources and sinks. *Tellus B*, *66*(1), 22486. doi: 10.3402/tellusb.v66.22486

Liu, J., Bowman, K. W., Schimel, D. S., Parazoo, N. C., Jiang, Z., Lee, M., ... Eldering, A. (2017). Contrasting carbon cycle responses of the tropical continents to the 2015–2016 el niño. *Science*, *358*(6360). doi: 10.1126/science.aam5690

Liu, Y. Y., Dorigo, W. A., Parinussa, R., de Jeu, R. A., Wagner, W., McCabe, M. F., ... Van Dijk, A. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens. Environ.*, *123*, 280–297.

Liu, Y. Y., Parinussa, R., Dorigo, W. A., De Jeu, R. A., Wagner, W., Van Dijk, A., ... Evans, J. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol. Earth Syst. Sc.*, *15*(2), 425–436.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., ... Ogawa, T. (2008). Worldwide measurements of atmospheric co2 and other trace gas species using commercial airlines. *Journal of Atmospheric and Oceanic Technology*, *25*(10), 1744–1754.

Machida, T., Matsueda, H., Sawa, Y., & Niwa, Y. (2018). *Atmospheric co_2 mole fraction data of CONTRAIL-CME, Ver.2017.1.0, center for global environmental research, NIES (reference date: 2018/11/28).* doi: 10.17595/20180208.001

Masarie, K., Peters, W., Jacobson, A., & Tans, P. (2014). Obspack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements. *Earth Syst. Sci. Data*, *6*(2), 375–384.

Menemenlis, D., Campin, J.-M., Heimbach, P., Hill, C., Lee, T., Nguyen, A., ... Zhang, H. (2008). Ecco2: High resolution global ocean and sea ice data synthesis. *Mercator Ocean Quarterly Newsletter*, *31*(October), 13–21.

Miller, S. M., & Michalak, A. M. (2019). The impact of improved satellite retrievals on estimates of biospheric carbon balance. *Atmos. Chem. Phys. Discuss.*, *2019*, 1–15. Retrieved from `https://www.atmos-chem-phys-discuss.net/acp-2019-382/` doi: 10.5194/acp-2019-382

Morino, I., Yokozeki, N., Matsuzaki, T., & Horikawa, M. (2017). *Tccon data from rikubetsu (jp), release ggg2014.r2.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/957` doi: 10.14291/tccon.ggg2014.rikubetsu01.r2

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., & Deng, F. (2013). Improving the temporal and spatial distribution of co2 emissions from global fossil fuel emission data sets. *J. Geophys. Res.-Atmos.*, *118*(2), 917–933.

Nassar, R., Sioris, C. E., Jones, D., & McConnell, J. C. (2014). Satellite observations of CO_2 from a highly elliptical orbit for studies of the Arctic and boreal carbon cycle. *J. Geophys. Res.-Atmos*, *119*(5), 2654–2673.

Nelson, R. R., & O'Dell, C. W. (2019). The impact of improved aerosol priors on near-infrared measurements of carbon dioxide. *Atmos. Meas. Tech.*, *12*(3), 1495–1512. Retrieved from `https://www.atmos-meas-tech.net/12/1495/2019/` doi: 10.5194/amt-12-1495-2019

Oda, T., & Maksyutov, S. (2011). A very high-resolution (1 km× 1 km) global fossil fuel co 2 emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.*, *11*(2), 543–556.

Oda, T., Maksyutov, S., & Andres, R. J. (2018). The open-source data inventory for anthropogenic co_2, version 2016 (odiac2016): a global monthly fossil fuel co_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth Syst. Sci. Data*, *10*(1), 87–107. Retrieved from `https://www.earth-syst-sci-data.net/10/87/2018/` doi: 10.5194/essd-10-87-2018

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., ... Wunch, D. (2012). The ACOS CO_2 retrieval algorithm – part 1: Description and validation against synthetic observations. *Atmos. Meas. Tech.*, *5*(1), 99–121. Retrieved from `http://www.atmos-meas-tech.net/5/99/2012/` doi:

10.5194/amt-5-99-2012

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., . . . Velazco, V. A. (2018). Improved retrievals of carbon dioxide from orbiting carbon observatory-2 with the version 8 acos algorithm. *Atmos. Meas. Tech.*, *11*(12), 6539–6576. Retrieved from https://www.atmos-meas-tech.net/11/6539/2018/ doi: 10.5194/amt-11-6539-2018

Olsen, S. C., & Randerson, J. T. (2004). Differences between surface and column atmospheric CO_2 and implications for carbon cycle research. *J. Geophys. Res.-Atmos.*, *109*(D2). (D02301) doi: 10.1029/2003JD003968

Parazoo, N. C., Denning, A. S., Kawa, S. R., Pawson, S., & Lokupitiya, R. (2012). CO_2 flux estimation errors associated with moist atmospheric processes. *Atmos. Chem. Phys.*, *12*(14), 6405–6416. doi: 10.5194/acp-12-6405-2012

Philip, S., Johnson, M. S., Potter, C., Genovesse, V., Baker, D. F., Haynes, K. D., . . . Poulter, B. (2019). Prior biosphere model impact on global terrestrial CO_2 fluxes estimated from oco-2 retrievals. *Atmos. Chem. Phys. Discuss.*, *2019*, 1–29. doi: 10.5194/acp-2018-1095

Polavarapu, S. M., Deng, F., Byrne, B., Jones, D. B. A., & Neish, M. (2018). A comparison of posterior atmospheric CO_2 adjustments obtained from in situ and gosat constrained flux inversions. *Atmos. Chem. Phys.*, *18*(16), 12011–12044. doi: 10.5194/acp-18-12011-2018

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

Randerson, J. T., Van Der Werf, G., Giglio, L., Collatz, G., & Kasibhatla, P. (2018). Global fire emissions database, version 4.1 (gfedv4). *ORNL DAAC*. doi: 10.3334/ORNLDAAC/1293

Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P., Koster, R. D., & De Lannoy, G. J. (2017). Assessment of MERRA-2 land surface hydrology estimates. *J. Climate*, *30*(8), 2937–2960.

Reichle, R. H., Koster, R. D., De Lannoy, G. J., Forman, B. A., Liu, Q., Mahanama, S. P., & Touré, A. (2011). Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, *24*(24), 6322–6338.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., . . . others (2014). Satellite-inferred european carbon sink larger than expected. *Atmos. Chem. Phys.*, *14*(24), 13739–13753.

Rödenbeck, C., Houweling, S., Gloor, M., & Heimann, M. (2003). CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmos. Chem. Phys.*, *3*(6), 1919–1964.

Saeki, T., Maksyutov, S., Saito, M., Valsala, V., Oda, T., RJ, A., . . . others (2013). Inverse modeling of CO_2 fluxes using gosat data and multi-year ground-based observations. *Sola*, *9*, 45–50.

Sasakawa, M., Machida, T., Tsuda, N., Arshinov, M., Davydov, D., Fofonov, A., & Krasnov, O. (2013). 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.-Atmos.*, *118*(16), 9489–9498. doi: 10.1002/jgrd.50755

Sasakawa, M., Shimoyama, K., Machida, T., Tsuda, N., Suto, H., Arshinov, M., . . . others (2010). Continuous measurements of methane from a tower network over Siberia. *Tellus B*, *62*(5), 403–416. doi: 10.1111/j.1600-0889.2010.00494.x

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., . . . others (2019). Quantifying the impact of atmospheric transport uncertainty on co2 surface flux estimates. *Global Biogeochem. Cy.*, *33*(4), 484–500.

Sellers, P., Randall, D., Collatz, G., Berry, J., Field, C., Dazlich, D., . . . Bounoua, L. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs.

part I: Model formulation. *J. Climate*, *9*(4), 676–705.

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., & Randall, D. A. (1996). A revised land surface parameterization (SiB2) for atmospheric GCMs. part II: The generation of global fields of terrestrial biophysical parameters from satellite data. *J. Climate*, *9*(4), 706–737.

Strong, K., Roche, S., Franklin, J. E., Mendonca, J., Lutsch, E., Weaver, D., ... Lindenmaier, R. (2017). *Tccon data from eureka (ca), release ggg2014.r2.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/970` doi: 10.14291/tccon.ggg2014.eureka01.r2

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., ... others (2015). Seasonal climatology of CO_2 across North America from aircraft measurements in the NOAA/ESRL Global Greenhouse Gas Reference Network. *J. Geophys. Res.-Atmos.*, *120*(10), 5155–5190.

Tans, P. P., Conway, T. J., & Nakazawa, T. (1989). Latitudinal distribution of the sources and sinks of atmospheric carbon dioxide derived from surface observations and an atmospheric transport model. *J. Geophys. Res.-Atmos.*, *94*(D4), 5151–5172.

Tramontana, G., Jung, M., Schwalm, C. R., Ichii, K., Camps-Valls, G., Ráduly, B., ... Papale, D. (2016). Predicting carbon dioxide and energy fluxes across global FLUXNET sites with regression algorithms. *Biogeosciences*, *13*(14), 4291–4313. doi: 10.5194/bg-13-4291-2016

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Kasibhatla, P. S., & Arellano Jr, A. F. (2006). Interannual variability in global biomass burning emissions from 1997 to 2004. *Atmos. Chem. Phys.*, *6*(11), 3423–3441.

Vidale, P., & Stöckli, R. (2005). Prognostic canopy air space solutions for land surface exchanges. *Theor. Appl. Climatol.*, *80*(2-4), 245–257.

Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., ... others (2012). Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals)*, *7*, 315–321.

Wang, J. S., Kawa, S. R., Collatz, G. J., Sasakawa, M., Gatti, L. V., Machida, T., ... Manyin, M. E. (2018). A global synthesis inversion analysis of recent variability in co_2 fluxes using gosat and in situ observations. *Atmos. Chem. Phys.*, *18*(15), 11097–11124. Retrieved from `https://www.atmos-chem-phys.net/18/11097/2018/` doi: 10.5194/acp-18-11097-2018

Warneke, T., Messerschmidt, J., Notholt, J., Weinzierl, C., Deutscher, N. M., Petri, C., & Grupe, P. (2017). *Tccon data from orléans (fr), release ggg2014.r0.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/283` doi: 10.14291/tccon.ggg2014.orleans01.r0/1149276

Wennberg, P. O., Roehl, C. M., Wunch, D., Toon, G. C., Blavier, J.-F., Washenfelder, R., ... Ayers, J. (2017). *Tccon data from park falls (us), release ggg2014.r1.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/295` doi: 10.14291/tccon.ggg2014.parkfalls01.r1

Wennberg, P. O., Wunch, D., Roehl, C. M., Blavier, J.-F., Toon, G. C., & Allen, N. T. (2017). *Tccon data from lamont (us), release ggg2014.r1.* CaltechDATA. Retrieved from `https://data.caltech.edu/records/279` doi: 10.14291/tccon.ggg2014.lamont01.r1/1255070

Wofsy, S. C. (2011). HIAPER pole-to-pole observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols. *Philos. T. R. Soc. A.*, *369*(1943), 2073–2086.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., ... Wennberg, P. O. (2011). The Total Carbon Column Observing Network. *Philos. T. Roy. Soc. A*, *369*(1943), 2087–2112. doi: 10.1098/rsta.2010.0240

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., ... Zondlo, M. A. (2010). Calibration of the total carbon column observing network using aircraft profile data. *Atmos. Meas. Tech.*, *3*(5), 1351–1362. doi: 10.5194/amt-3-1351-2010

Yokota, T., Yoshida, Y., Eguchi, N., Ota, Y., Tanaka, T., Watanabe, H., & Maksyutov, S. (2009). Global concentrations of co2 and ch4 retrieved from gosat: First preliminary results. *Sola*, *5*, 160–163.

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., ... others (2013). Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data.

Yu, K., Keller, C. A., Jacob, D. J., Molod, A. M., Eastham, S. D., & Long, M. S. (2018). Errors and improvements in the use of archived meteorological data for chemical transport modeling: an analysis using GEOS-Chem v11-01 driven by GEOS-5 meteorology. *Geosci. Model Dev.*, *11*(1), 305–319. doi: 10.5194/gmd-11-305-2018

Supporting Information for "Improved constraints on northern extratropical CO_2 fluxes obtained by combining surface-based and space-based atmospheric CO_2 measurements"

B. Byrne[1], J. Liu[2], M. Lee[2], I. Baker[3], K. W. Bowman[2], N. M. Deutscher[4],

D. G. Feist[5], D. W. T. Griffith[4], L. T. Iraci[6], M. Kiel[2], J. S. Kimball[7],

C. E. Miller[2], I. Morino[8], N. C. Parazoo[2], C. Petri[9], C. M. Roehl[10],

M. K. Sha[11], K. Strong[12], V. A. Velazco[4], P. O. Wennberg[10,13], D. Wunch[12]

[1]NASA Postdoctoral Program Fellow, Jet Propulsion Laboratory, California Institute of Technology, CA, USA

[2]Jet Propulsion Laboratory, California Institute of Technology, CA, USA

[3]Atmospheric Science Department, Colorado State University, Fort Collins, CO, USA

[4]Centre for Atmospheric Chemistry, University of Wollongong, Wollongong, Australia

[5]Max Planck Institute for Biogeochemistry, Jena, Germany

[6]Atmospheric Science Branch, NASA Ames Research Center, Moffett Field, CA 94035, USA

[7]Numerical Terradynamic Simulation Group, W.A. Franke College of Forestry & Conservation, The University of Montana, Missoula, MT 59812, USA

[8]Satellite Observation Center, Center for Global Environmental Research, National Institute for Environmental Studies (NIES), 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan

[9]Institute of Environmental Physics, University of Bremen, Bremen, Germany

[10]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

[11]Royal Belgian Institute for Space Aeronomy (BIRA-IASB), Brussels, Belgium

[12]Department of Physics, University of Toronto, Toronto, Ontario, Canada

[13]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA, USA

November 8, 2019, 1:42am

Contents of this file

1. Figures S1 to S7
2. Tables S1 to S2

Corresponding author: B. Byrne, Jet Propulsion Laboratory M/S 233-200, 4800 Oak Grove Drive, Pasadena, CA 91109. (brendan.k.byrne@jpl.nasa.gov)

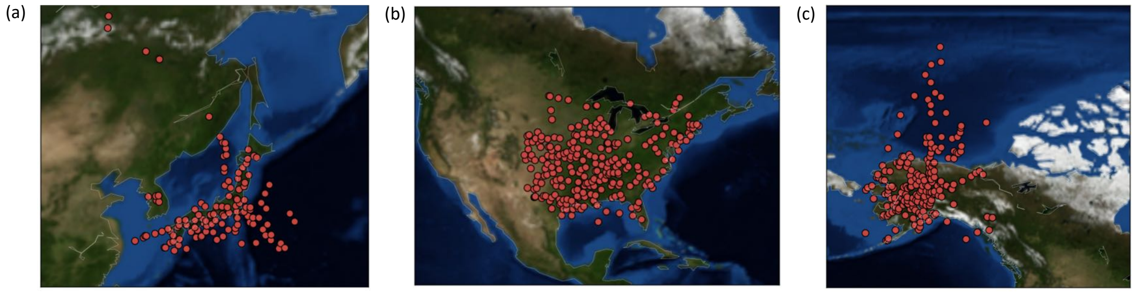

Figure S1. Locations of aircraft observations used in this study for (a) East Asia, (b) North America, and (c) Alaska/Arctic.

Table S1. Mean and standard deviation (std) of data–model mismatch between each flux inversion and aircraft-based CO_2 observations over East Asia, North America, and Alaska/Arctic. Posterior-simulated-CO_2 was calculated at 4° × 5° spatial resolution.

Region		East Asia		North America		Alaska/Arctic	
data set	prior NEE	mean (ppm)	std (ppm)	mean (ppm)	std (ppm)	mean (ppm)	std (ppm)
prior	SiB3	-0.06	0.85	0.08	0.97	-0.84	1.61
	CASA	-0.01	0.76	0.26	0.56	-0.59	1.36
	FLUXCOM	1.18	0.70	1.54	0.57	1.24	1.00
	Mean NEE	0.37	0.57	0.63	0.54	-0.06	1.16
TCCON	SiB3	0.16	0.46	0.33	0.43	-0.10	0.86
	CASA	0.33	0.74	0.65	0.57	-0.02	1.30
	FLUXCOM	0.42	0.45	0.42	0.45	-0.02	1.18
	Mean NEE	0.30	0.42	0.43	0.47	-0.05	1.05
surface-only	SiB3	0.01	0.44	0.34	0.35	-0.06	0.80
	CASA	0.13	0.71	0.48	0.50	-0.14	1.22
	FLUXCOM	0.22	0.60	0.46	0.33	-0.01	0.88
	Mean NEE	0.12	0.43	0.43	0.31	-0.07	0.93
GOSAT-only	SiB3	0.25	0.41	0.49	0.37	-0.06	0.76
	CASA	0.14	0.36	0.43	0.36	-0.17	0.81
	FLUXCOM	0.23	0.44	0.50	0.33	0.03	0.89
	Mean NEE	0.21	0.33	0.47	0.32	-0.06	0.79
GOSAT +surface +TCCON	SiB3	0.18	0.35	0.34	0.31	-0.7	0.75
	CASA	0.15	0.39	0.42	0.36	-0.03	0.89
	FLUXCOM	0.16	0.38	0.39	0.32	0.00	0.93
	Mean NEE	0.16	0.31	0.38	0.32	-0.03	0.84

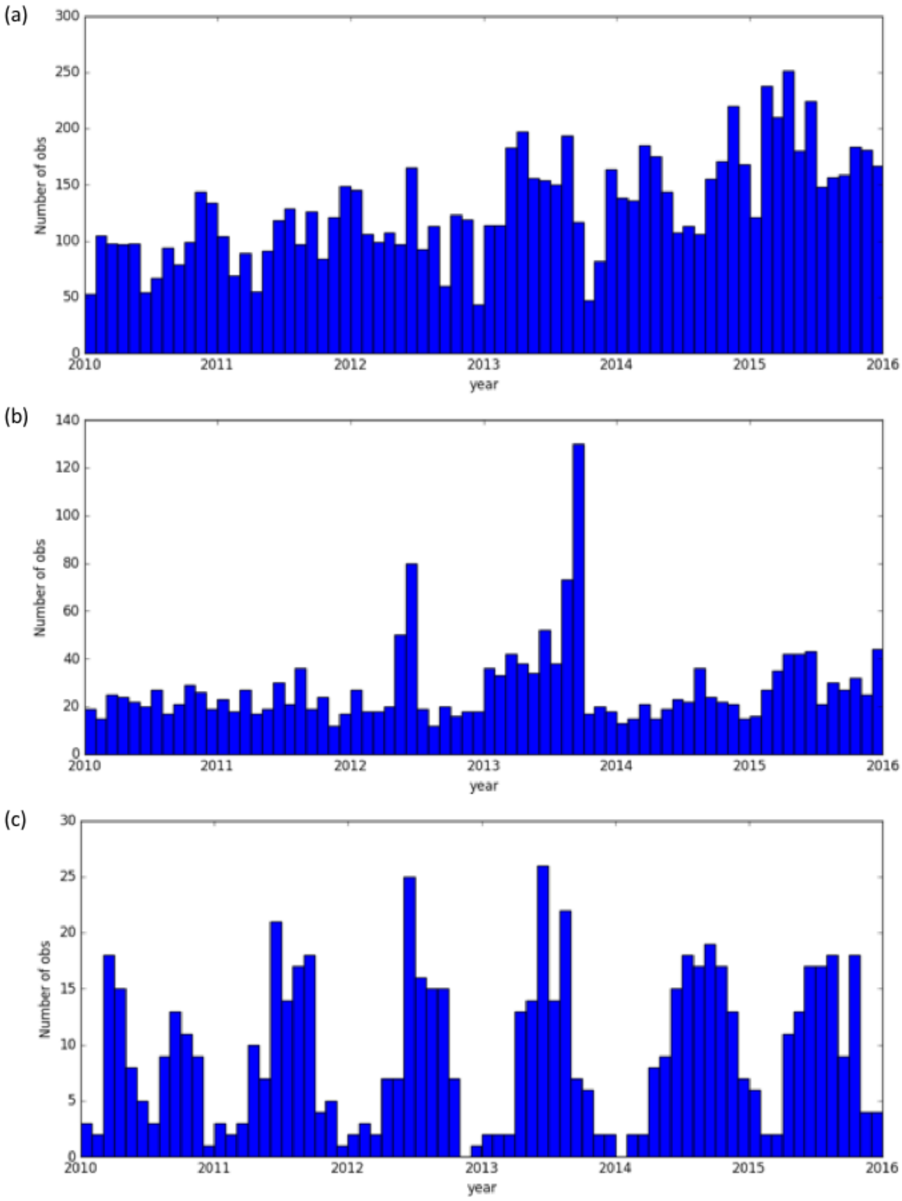

Figure S2. Number of hourly-mean aircraft measurements between 3–8 km altitude above sea level per month for (a) East Asia, (b) North America, and (c) Alaska/Arctic.

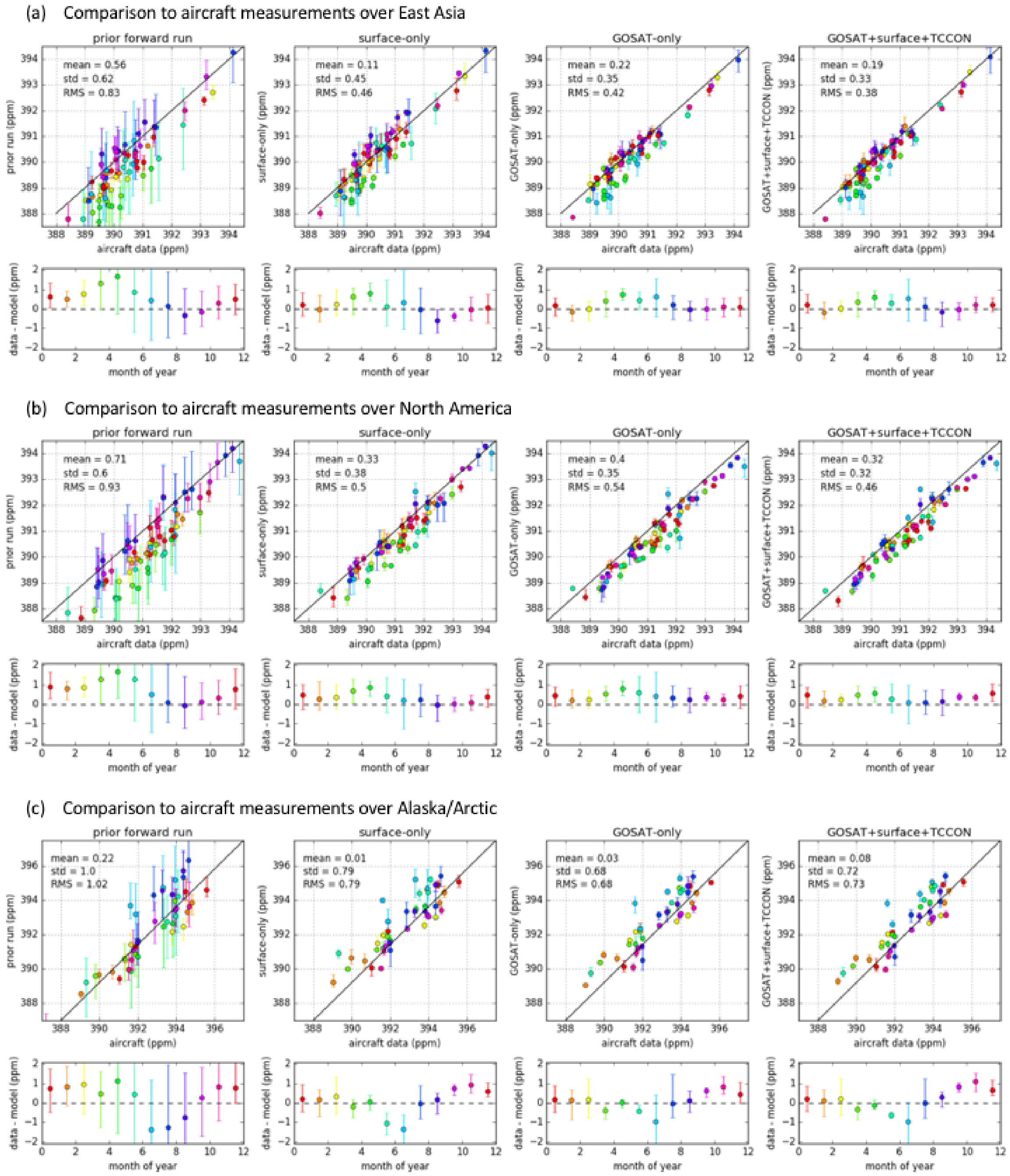

Figure S3. Same as Fig. 3 but at $2° \times 2.5°$ spatial resolution (except for TCCON). Comparison of monthly mean measured and simulated aircraft-based CO_2 for (a) East Asia, (b) North America, and (c) Alaska/Arctic. For each region, the mismatch for (left to right) prior, surface-only, GOSAT-only, and GOSAT+surface+TCCON simulated CO_2 are shown. The top panel shows a scatter plot of the simulated aircraft-based CO_2 against the measured aircraft-based CO_2, and the error bars indicate the spread in posterior NEE. The lower panel shows the mean data–model mismatch for each month, with error bars showing the range of monthly mean mismatched over the six-years and inversion set-ups. Colors correspond to the month of year.

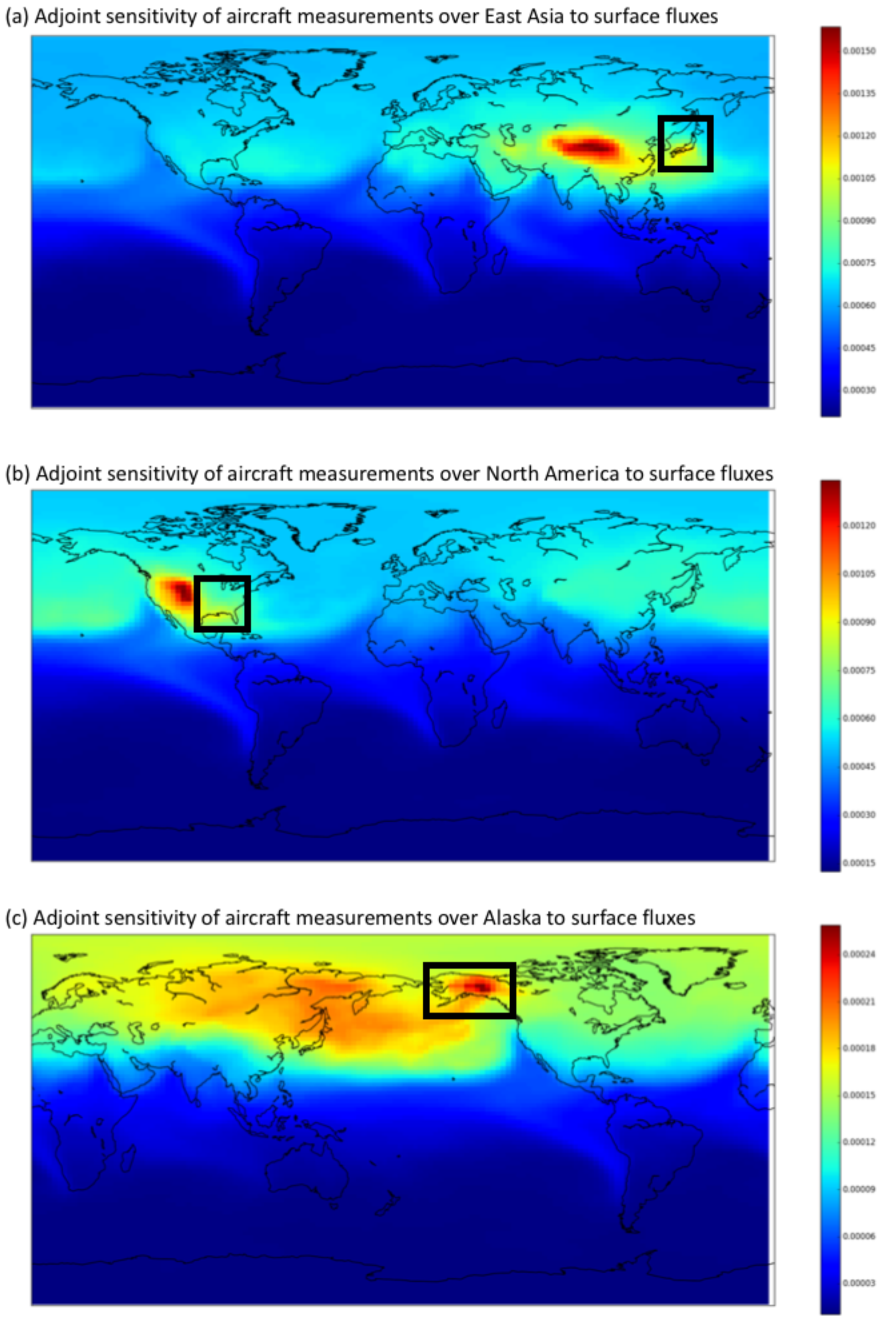

Figure S4. Adjoint sensitivity of aircraft-based CO_2 measurements to surface fluxes for measurements over (a) East Asia, (b) North America, and (c) Alaska/Arctic. Black boxes show the location of aircraft-based CO_2 measurements.

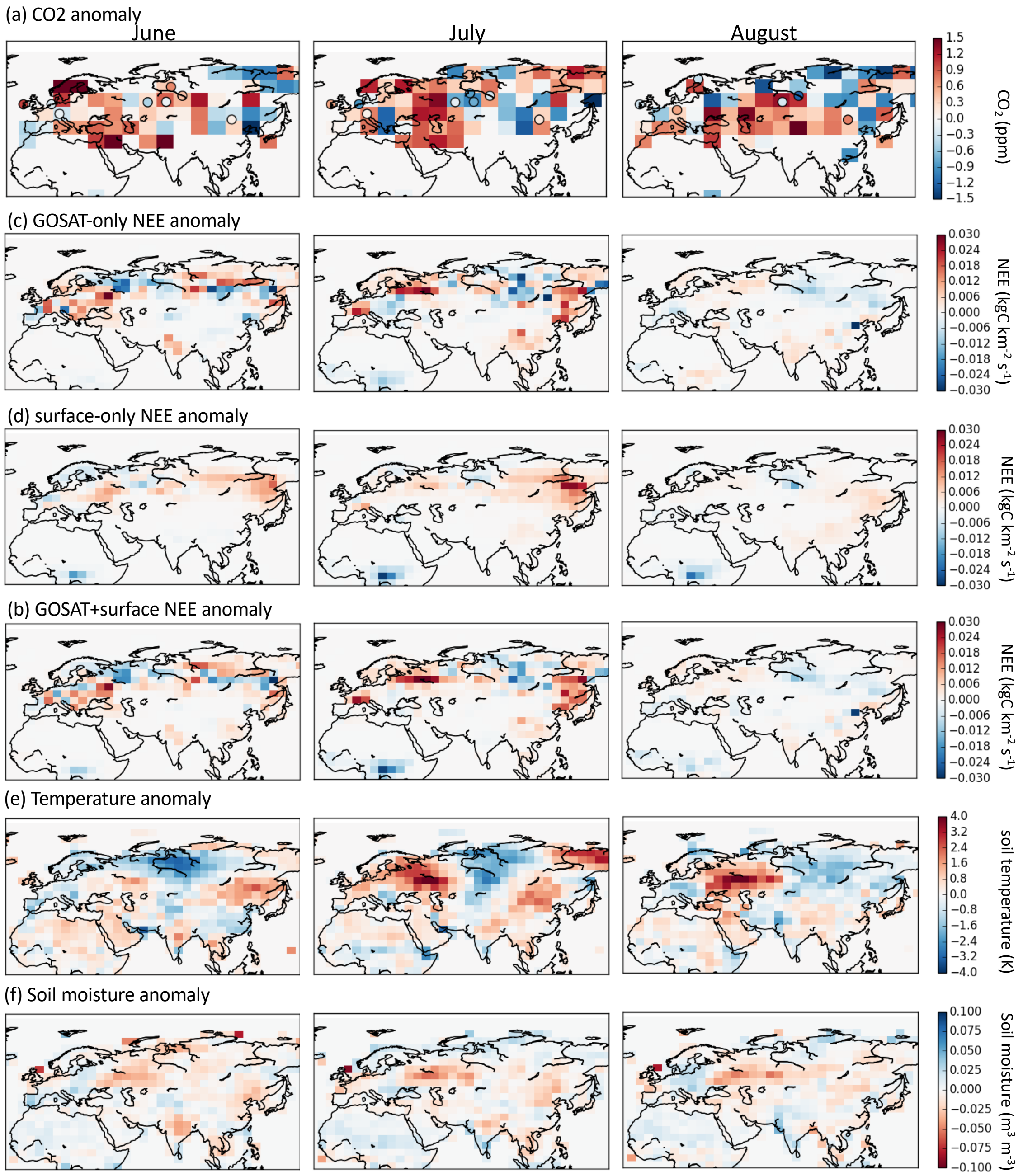

Figure S5. Same as Fig. 8 but for Eurasia during (left-to-right) May, June, July and August of 2010. Monthly anomalies in (a) GOSAT X_{CO_2} (ppm, 4° × 5° grid cells) and surface site CO_2 (ppm divided by four, circles), (b) GOSAT-only posterior NEE, (c) surface-only posterior NEE, (d) GOSAT+surface posterior NEE, (e) MERRA-2 soil temperature anomalies (K), and (f) ESA CCI soil moisture.

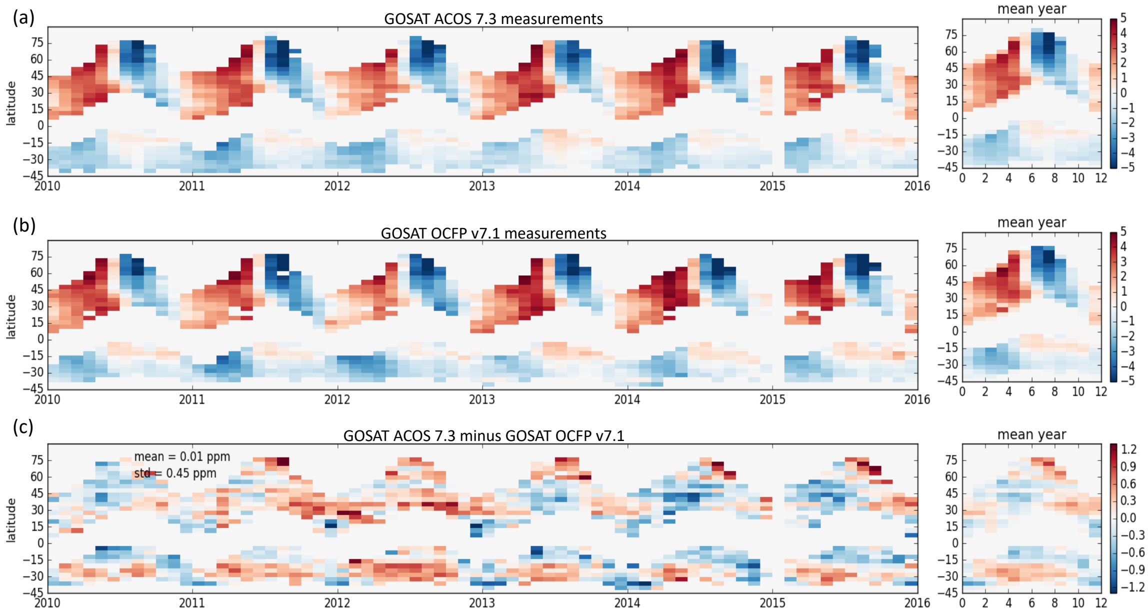

Figure S6. Detrended zonal-monthly mean high-gain nadir GOSAT X_{CO_2} retrieved by (a) ACOS 7.3 and (b) OCFP v7.1. (c) Difference in X_{CO_2} between the two retrieval algorithms.

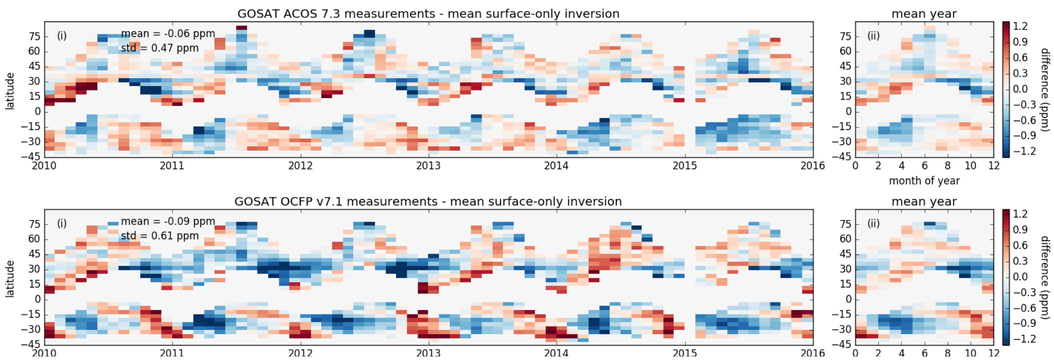

Figure S7. Data–model mismatch of the (a) ACOS 7.3 and (b) OCFP v7.1 GOSAT high-gain nadir X_{CO_2} measurements as a function of latitude and time for the surface-only flux inversion.

Table S2. Mean and standard deviation (std) of data–model mismatch between each flux inversion and aircraft-based CO_2 observations over East Asia, North America, and Alaska/Arctic. Posterior-simulated-CO_2 was calculated at $2^\circ \times 2.5^\circ$ spatial resolution.

Region		East Asia		North America		Alaska/Arctic	
data set	prior NEE	mean (ppm)	std (ppm)	mean (ppm)	std (ppm)	mean (ppm)	std (ppm)
4prior	SiB3	0.57	0.94	0.56	1.03	0.01	1.56
	CASA	-0.05	0.73	0.18	0.57	-0.54	1.20
	FLUXCOM	1.16	0.75	1.39	0.62	1.19	0.90
	Mean NEE	0.56	0.62	0.71	0.60	0.22	1.00
surface-only	SiB3	0.01	0.44	0.26	0.40	0.03	0.73
	CASA	0.11	0.69	0.38	0.57	-0.06	1.04
	FLUXCOM	0.22	0.62	0.35	0.39	0.06	0.79
	Mean NEE	0.11	0.45	0.33	0.38	0.01	0.79
GOSAT-only	SiB3	0.25	0.38	0.42	0.38	0.03	0.65
	CASA	0.18	0.39	0.37	0.39	-0.07	0.72
	FLUXCOM	0.24	0.46	0.42	0.36	0.14	0.75
	Mean NEE	0.22	0.35	0.40	0.35	0.03	0.68
GOSAT	SiB3	0.20	0.37	0.28	0.33	0.06	0.66
+surface	CASA	0.15	0.40	0.33	0.39	0.04	0.78
+TCCON	FLUXCOM	0.22	0.38	0.36	0.32	0.15	0.78
	Mean NEE	0.19	0.33	0.32	0.32	0.08	0.72