Regional-scale, sector-specific evaluation of global CO2 inversion models using aircraft data from the ACT-America project

Brian Gaudet[1], K J Davis[1], S Pal[2], A R Jacobson[3], A Schuh[4], T Lauvaux[5], S Feng[1], and E V Browell[6]

[1]The Pennsylvania State University
[2]Texas Tech University
[3]CIRES
[4]CIRA
[5]Laboratoire des Sciences du Climat et de l'Environnement
[6]STARSS-III Affiliate

November 23, 2022

Abstract

We use 148 airborne vertical profiles of CO2 for frontal cases from the summer 2016 Atmospheric Carbon and Transport-America (ACT-America) campaign to evaluate the skill of ten global CO2 in situ inversion models from the version 7 Orbiting Carbon Observatory-2 (OCO-2) Model Intercomparison Project (MIP). Model errors (model posterior-observed CO2 dry air mole fractions) were categorized by region (Mid-Atlantic, Midwest, and South), frontal sector (warm or cold), and transport model (predominantly Tracer Model 5 (TM5) and Goddard Earth Observing System-Chemistry (GEOS-Chem)). All inversions assimilated the same CO2 observations. Overall, the median inversion profiles reproduce the general structures of the observations (enhanced / depleted low-level CO2 in warm / cold sectors), but 1) they underestimate the magnitude of the warm / cold sector mole fraction difference, and 2) the spread among individual inversions can be quite large (> 5 ppm). Uniquely in the Mid-Atlantic, inversion biases segregated according to atmospheric transport model, where TM5 inversions biases were-3 to-4 ppm in warm sectors, while those of GEOS-Chem were +2 to +3 ppm in cold sectors. The large spread among the mean posterior CO2 profiles is not explained by the different atmospheric transport models. These results show that the inversion systems themselves are the dominant cause of this spread, and that the aircraft campaign data are clearly able to identify these large biases. Future controlled experiments should identify which inversions best reproduce midlatitude CO2 mole fractions, and how inversion system components are linked to system performance.

Hosted file

`agusupporting-information_word_template.docx` available at https://authorea.com/users/541417/articles/600612-regional-scale-sector-specific-evaluation-of-global-co2-inversion-models-using-aircraft-data-from-the-act-america-project

Regional-scale, sector-specific evaluation of global CO_2 inversion models using aircraft data from the ACT-America project

B. J. Gaudet[1,2*], K. J. Davis[1,3], S. Pal[4], A.R. Jacobson[5], A. Schuh[6], T. Lauvaux[7], S. Feng[1], and E.V. Browell[8]

[1]Department of Meteorology and Atmospheric Science, The Pennsylvania State University.

[2]Pacific Northwest National Laboratory.

[3]Earth and Environmental Systems Institute, The Pennsylvania State University

[4]Department of Geosciences, Atmospheric Science Division, Texas Tech University.

[5]CIRES, University of Colorado Boulder and NOAA Earth System Research Laboratory, Global Monitoring Laboratory.

[6]CIRA, Colorado State University.

[7]Laboratoire des Sciences du Climat et de l'Environnement.

[8]STARSS-III Affiliate, NASA Langley Research Center.

Corresponding author: Brian J. Gaudet (brian.gaudet@pnnl.gov)

[*] Corresponding author address: Brian J. Gaudet, Earth Systems Analysis and Modeling Group, Pacific Northwest National Laboratory, PO Box 999, M/S K9-30, Richland, WA 99352.

Key Points:

- Global inversions capture observed CO_2 variations across North America for summer 2016, but underestimate warm / cold sector differences.
- Inversion variability appears to be mainly driven by differences in inversion system design and/or prior fluxes, not transport model.
- Uniquely in the Mid-Atlantic, transport-dependent CO_2 biases exist, present in both warm and cold sectors.

Abstract

We use 148 airborne vertical profiles of CO_2 for frontal cases from the summer 2016 Atmospheric Carbon and Transport – America (ACT-America) campaign to evaluate the skill of ten global CO_2 *in situ* inversion models from the version 7 Orbiting Carbon Observatory – 2 (OCO-2) Model Intercomparison Project (MIP). Model errors (model posterior – observed CO_2 dry air mole fractions) were categorized by region (Mid-Atlantic, Midwest, and South), frontal sector (warm or cold), and transport model (predominantly Tracer Model 5 (TM5) and Goddard Earth Observing System – Chemistry (GEOS-Chem)). All inversions assimilated the same CO_2 observations. Overall, the median inversion profiles reproduce the general structures of the observations (enhanced / depleted low-level CO_2 in warm / cold sectors), but 1) they underestimate the magnitude of the warm / cold sector mole fraction difference, and 2) the spread among individual inversions can be quite large (> 5 ppm). Uniquely in the Mid-Atlantic, inversion biases segregated according to atmospheric transport model, where TM5 inversions biases were -3 to -4 ppm in warm sectors, while those of GEOS-Chem were +2 to +3 ppm in cold sectors. The large spread among the mean posterior CO_2 profiles is not explained by the different atmospheric transport models. These results show that the inversion systems themselves are the dominant cause of this spread, and that the aircraft campaign data are clearly able to identify these large biases. Future controlled experiments should identify which inversions best reproduce midlatitude CO_2 mole fractions, and how inversion system components are linked to system performance.

1 Introduction

To mitigate the effects of climate change due to greenhouse gases (GHGs), it is vital that we understand the exchange of CO_2 between the atmosphere, biosphere, and ocean, and how these exchanges might vary with time. Previous studies have shown that the land biosphere is a substantial sink of atmospheric CO_2 at present (e.g., Tans et al., 1990; Denning et al., 1995; Gurney et al., 2004; Ballantyne et al. 2012), but much remains unknown about its precise magnitude and long-term trajectory, what accounts for its variability, and to what extent it is tropical or extratropical (Schimel et al., 2015; Crowell et al., 2019). In part, this lack of understanding is due to a gap in observational coverage at regional-to-continental scales (spatially, approximately 10^3 – 10^4 km). While eddy flux towers can provide measurements of surface CO_2 fluxes at local scales (Luyaessert et al., 2007), and hemispheric-to-global fluxes can be constrained by mole fraction measurements at remote sites (Ciais et al., 2019), at intermediate scales neither method is completely satisfactory. Measurements of CO_2 mole fraction from *in situ* stations do not have a straightforward relationship to surface fluxes on intermediate scales (Stephens et al., 2007), and these stations are predominantly located in the mid-latitudes, limiting the assessment of surface fluxes in tropical and high-latitude regions (Peylin et al., 2013).

The current state-of-the-art method of generating a full atmospheric analysis of CO_2 mole fractions and surface fluxes on regional-to-global scales is through the use of global flux inversion models (Enting et al., 1995; Bousquet et al., 2000; Peylin et al., 2013; Basu et al., 2018). These Bayesian optimization systems use global atmospheric transport models to simulate mole fraction fields from a given prior distribution of surface fluxes, and then "invert" the transport to find the optimal surface fluxes that minimize errors between observed mole fractions and those that result from forward simulations of those optimal surface fluxes ('posterior CO_2'; Tarantola, 2005). The posterior mole fraction field provides an analysis of unobserved CO_2 distributions in synoptic weather systems down to the resolution limit of these models, generally a few degrees of latitude / longitude (e.g. Peters et al. 2007). The accuracy of these global inversions, however, is dependent on choice of data inversion procedure, prior flux model, and transport model used within the global inversions, any of which can vary considerably among different inversion groups (Peylin et al., 2013). In particular, this often manifests itself as an inability to associate errors in modeled CO_2 uniquely with transport or surface flux errors. The absence of sufficiently dense independent data networks hinders the ability to support or refute particular inversion results.

Tropospheric profiles of GHGs have been obtained for the last two decades at approximately biweekly intervals by the National Oceanographic and Atmospheric Administration (NOAA) light aircraft profiler network at over 22 sites in North America, and have provided independent assessment of the skill of global inversions throughout the troposphere (Sweeney et al., 2015). These data have been used to show evidence of increased CO_2 depletion in the boreal growing season from west-to-east across the North American continent, consistent with a surface biogenic sink and a prevailing eastward motion of air masses, though some of the depletion was attributed to Eurasian sinks upstream (Sweeney et al., 2015; Lan et al., 2017). The CarbonTracker global inversion system, version CT2015 (Peters et al., 2007, with updates documented at http://carbontracker.noaa.gov) was shown by Lan et al. (2017) to have horizontal gradients of column-averaged CO_2 (XCO_2) that compare well with the available data. Stephens et al. (2007) used the profiles to evaluate the seasonal vertical CO_2 gradients of 12 global inversions from the Transcom 3 intercomparison experiment (Gurney et al., 2004; Baker et al., 2006), and found that the inversions tended to underestimate the observed positive vertical gradient during the boreal growing season, which was attributed to overestimation of vertical mixing. It was shown on these timescales that overestimated vertical mixing in an inversion that assimilates near-surface mole fraction data led to an overestimated net surface flux of CO_2 , though the correlation was weaker during the growing season. The recent study of Gaubert et al. (2019) found better agreement between inversion vertical CO_2 gradients and oceanic observations in the annual mean, but biased vertical gradients were still present when just boreal summer was considered.

The spatial and temporal scale of the NOAA long-term aircraft profiles are too coarse to resolve structures found in individual synoptic weather systems, which progress over the continent on a time scale of days, and have airstream sectors (warm conveyor belt, cold conveyor belt, dry intrusion) on scales of hundreds to thousands of km. Furthermore, as noted by Sweeney et al. (2015), the profiles tend to preferentially sample fair-weather conditions. While over a long time period these observations help constrain CO_2 fluxes and transport, they do not adequately sample the structures of synoptic weather systems that play an important role in vertical and latitudinal transport of CO_2 (Schuh et al, 2019). More recently, field campaigns using more extensive sampling of GHG with aircraft have occurred, such as the Atmospheric Tomography Mission (ATom; Wofsy et al. 2018) and the O_2 / N_2 Ratio and CO_2 Airborne Southern Ocean Study (ORCAS; Stephens et al. 2018), but these generally did not target the North American continent. Tower-based observations (Andrews et al., 2014), while continuous in time are spatially sparse and limited to the planetary boundary layer (PBL). Model-data comparisons using tower-based, *in situ* CO_2 observations have shown large model-data differences (Diaz-Isaac et al, 2014; 2018) but have not yielded direct insight into the representation of weather systems in atmospheric CO_2 simulations.

Recently, dedicated satellites have been launched that can infer XCO_2 from shortwave infrared observations, including the Greenhouse gases Observing SATellite (GOSAT: Kuze et al., 2009), and the Observing Carbon Observatory -- 2 (OCO-2) satellite platform launched by NASA in 2014 (Crisp 2005; Eldering et al., 2017). These programs have the goals of increasing global data and coverage and reducing observational dependence on vertical mixing. To assess the sensitivity of the global inversions to OCO-2 data as well as to the different choices of inversion procedure mentioned above, NASA organized the OCO-2 Model Intercomparison Project (OCO-2 MIP) including 11 global inversion groups (Crowell et al., 2019). As part of the OCO-2 MIP, a set of standardized numerical experiments with varying transport models, optimization techniques, and prior surface fluxes were performed with different sets of common assimilated data; one set of experiments assimilated OCO-2 retrievals, while another set only assimilated standardized *in situ* observations (henceforth the IS inversions). Among other findings, they found generally small, but positive (< about +1 ppm) biases among the IS inversions relative to independent aircraft observations in the northern extratropics. However, overall variability among models constrained with IS data remained high, and the independent aircraft datasets were deemed too sparse to be able to make more specific assessments of the skill of different inversion techniques.

Schuh et al. (2019) noted systematic differences within the OCO-2 MIP between inversions that used the Goddard Earth Observing System – Chemistry (GEOS-Chem; Bey et al., 2001) transport model, and those that used the Tracer Model 5 (TM5; Krol et al., 2015). Poleward of 45 N, the GEOS-Chem IS inversions had reduced growing season surface uptake, and overall seasonal flux amplitude, compared with the TM5 inversions, although large variability among the full inversion systems precluded strict statistical significance. They then performed controlled experiments, running forward simulations with the same surface fluxes (from CarbonTracker CT2016) for both TM5 and GEOS-Chem. These confirmed that, relative to TM5, GEOS-Chem had a tendency to ‘trap’ surface flux signals near the surface and advect them poleward; thus GEOS-Chem would require reduced seasonal cycles of local CO_2 surface fluxes to match the same set of near-surface mole fraction measurements. Outstanding questions include which transport model is closer to observations, and whether these characteristics are due to differences in vertical or horizontal mixing (e.g., GEOS-Chem could have greater meridional transport of CO_2 in the midlatitude Ferrel cells (Peixoto and Oort, 1992; Pauluis et al., 2009; Parazoo et al., 2011), increasing relative horizontal mixing but reducing relative vertical mixing of CO_2). TM5 vs. GEOS-Chem forward-transport experiments with the anthropogenic tracer SF_6 in Schuh et al. (2019) suggested that TM5 transport produced better agreement with marine near-surface observations, but transport model skill above the boundary layer was not known.

To increase our understanding of GHG mole fractions and fluxes over North America, the Atmospheric Carbon Transport -- America (ACT-America) NASA Earth Venture Suborbital 2 (EVS-2) mission (Davis et al., 2018; Baier et al., 2019; Pal et al. 2020) is using a combination of aircraft, satellite, and tower-based observational platforms, including a series of aircraft field campaigns that cover three focus regions (the U.S. Mid-Atlantic, Midwest, and South) and all four seasons, beginning with summer 2016. During each six-week campaign, flight plans for each individual flight day were designed based on forecast meteorology for either investigating multiple sectors of frontal weather systems, sampling large-scale mole fractions over fair weather boundary layers, or providing under-flights of OCO-2 passages to help evaluate its retrievals. It was hypothesized that the synoptic-scale variability of CO_2 on frontal days would have strong sensitivity to atmospheric transport, and that these data could be used to evaluate transport models and determine the optimal ones to use within flux inversion systems.

Pal et al. (2020) presents an observational analysis of the aircraft *in situ* CO_2 and CH_4 mole fractions from the frontal cases of the summer 2016 ACT-America campaign, classifying data as either 'warm sector' or 'cold sector', and within a sector into PBL, lower free troposphere, and upper free troposphere. In the PBL they found systematically greater CO_2 mole fractions in warm sectors than cold sectors (by 5-30 ppm), and, on average, warm / cold sector mole fractions were enhanced / depleted relative to the free troposphere. (Henceforth, 'enhanced' and 'depleted' mean higher / lower mole fractions relative to the free troposphere.) Free troposphere warm sectors also had enhanced CO_2 relative to cold sectors, but by reduced magnitude (5 ppm or less) compared to the PBL. Chen et al. (2019) used the vertical profile data from the summer 2016 and winter 2017 campaigns to evaluate the CO_2 mole fractions in a pair of global inversions available during the period -- the ECMWF-based Copernicus Atmosphere Monitoring Service (CAMS) (Agusti-Panareda et al., 2014), and CarbonTracker Near-Realtime (CT-NRT), a version of the CarbonTracker system that uses some climatologically-based flux priors and a subset of observations to deliver a product with reduced processing time. They found the two inversions agreed reasonably well with the independent ACT-America data. However, substantial low biases for CT-NRT were noted in summer 2016 over the MidAtlantic region and in winter 2017 over the Midwest; meanwhile, CAMS showed a positive low-level bias in the MidAtlantic during summer 2016, and a positive bias throughout the column in winter 2017. The uncertainty as inferred from the model spread was deemed comparable to the model error.

In this study, we also use *in situ* CO_2 from vertical profiles during the frontal cases from the summer 2016 ACT-America campaign, but here we apply the data to an assessment of the whole suite of inversion models participating in the version 7 OCO-2 MIP study. We analyze the structures and model errors of the profiles as a function of region, meteorological sector, and transport model. We use only flight segments corresponding to vertical profiles because they provide information about atmospheric vertical structure with a minimum of temporal and spatial variability. In particular, this allows the comparison of modeled and observed CO_2 mole fraction as a function of sector without the complication of mismatches between the model and observed frontal position. For simplicity we only examine the OCO-2 MIP inversions that assimilated a prescribed suite of *in situ* data (the IS experiment), with a special focus on the potential impact of transport model, which in the OCO-2 MIP was predominantly either TM5 or GEOS-Chem. In section 2 we provide more background on the data and model simulations and describe our experimental procedure. Section 3 presents the resulting statistics of these vertical profiles, while sections 4 and 5 include contain overall discussion and conclusions from the work.

2 Data and Methodology

2.1 ACT-America aircraft measurements

The six-week summer 2016 ACT-America campaign took place from 15 July – 31 August 2016. For three successive two-week periods, the base of operations for the C-130 and Kingair B-200 aircraft was at NASA Langley Research Center / Wallops Flight Facility (Mid-Atlantic region), Lincoln, NE (Midwest region), and Shreveport, LA (South region). A total of 25 research flight days occurred – 7 in the Mid-Atlantic, 9 in the Midwest, and 9 in the South. We analyzed only the frontal case flight days from each region (for a list of case days, see the Supplemental Information, Table S1). Flights were conducted in midday or afternoon hours to focus on well-mixed PBL conditions. We used CO_2 mole fractions derived from the Picarro sensors on board each aircraft. More information on the data used can be found in Davis et al. (2018) and Pal et al. (2020).

The aircraft data from the summer 2016 ACT-America campaign was put into Obspack format (see https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/OCO2_insitu/ for details). ACT-America data are publicly available from the NASA archive (https://www-air.larc.nasa.gov/cgi-bin/ArcView/actamerica.2016) and the ORNL Distributed Active Archive Center (Yang et al., 2018). Profiles identified for each ACT-America campaign are also available from ORNL (Pal, 2019).

2.2 OCO-2 MIP Suite

We used 10 different models from the version 7 OCO-2 MIP suite in this study, as listed in Table 1. All but one are global CO_2 flux inversion models (i.e., they optimize CO_2 surface fluxes on the basis of assimilated CO_2 observations). Three inversions use the TM5 transport model, which uses ERA-Interim analyses (Berrisford et al., 2011) as a meteorological driver, while four use the GEOS-Chem transport model, based on either the GEOS 'forward processing' (GEOS-FP) system or the 2nd version of the Modern Era Retrospective analysis for Research and Applications (MERRA-2) (Bosilovich et al., 2015) driver data. Of the remaining models, one ('GEOS Model') is not an inversion, but uses the parent transport model (GEOS-5) of GEOS-Chem (Reinecker et al., 2008). 'CAMS' is an inversion that uses the Laboratoire de Météorologie Dynamique – Zoom, version 3 (LMDz3) transport model (Chevalier et al. 2005), which makes use of ERA-Interim, and so should have similar transport to the TM5 family. The PCTM transport model also makes use of MERRA-2 driver data, like the GEOS-Chem members, but differs mainly in its modeling of vertical transport. Four variants of this model with different oceanic fluxes participated in the MIP, but since all were similar over our regions of interest, only the variant using the NASA Ocean Biogeochemical Model was used in this study. Other details about the MIP can be found in Crowell et al. (2019).

CO_2 mole fractions were extracted from all participating models in Obspack format at the times and locations corresponding to the ACT-America aircraft observations (also available at https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/OCO2_insitu /).

2.3 Method

We used 148 flight segments identified by Pal et al. (2020) as either warm sector or cold sector vertical profiles for analysis. Each frontal day contained both warm sector and cold sector profiles as noted in Table S1. We performed the analysis for each combination of two sectors and three regions, where for simplicity we assumed that the sector classification of each modeled profile was the same as its corresponding observation.

We binned both modeled and observed CO_2 mole fraction into 250-m bins of altitude above ground level (AGL), derived by taking the geopotential height of each model / observation pair and subtracting the average ground elevation for the profile as reported in the NASA ORNL DAAC netCDF files. We performed the analysis for bins greater than 250 m AGL, with the exception that we excluded the 250-500 m AGL bin for times before 1700 UTC, to exclude boundary layers that might not be well mixed. We also truncated the profile analysis when bins contained two or fewer sounding points.

Table 1. Global inversions / models participating in OCO-2 MIP (adapted from Crowell et al. (2019)).

Name	Transport	Driver	Method	Resolution (deg)	Reference
CT-NRT	TM5	ERA-Interim	EnKF	2x3 (1x1 N. Amer.)	Peters et al., 2007 (with online updates)
OU	TM5	ERA-Interim	4DVar	4x6	Crowell et al., 2018
CSU	GEOS-Chem	MERRA-2	Bayesian	1x1	Schuh et al., 2010
CMS-Flux	GEOS-Chem	GEOS-FP	4DVar	4x5	Liu et al., 2014
TM5-4DVar	TM5	ERA-Interim	4DVar	2x3	Basu et al., 2013
GEOS Model	GEOS-5	Inline	--	0.3125 x 0.25	Rienecker et al., 2008
CAMS	LMDZ3	ERA-Interim	Variational	1.875 x 3.75	Chevalier et al., 2005
U Toronto	GEOS-Chem	GEOS-FP	4DVar	4x5	Deng and Chen, 2011
U Edinburgh	GEOS-Chem	GEOS-FP	EnKF	4x5	Feng et al., 2016
PCTM	PCTM	MERRA-2		6.7x6.7	Baker et al., 2010

We estimated the variability of the mean observed mole fraction of each height bin with the conventional standard error of the mean, $\sigma/\sqrt{N}$, where σ is the standard deviation of all five-second observations in each height bin. Choosing N to be the number of observations within each height bin, however, is strictly appropriate for independent observations. In reality, the mole fractions within a particular vertical profile are highly autocorrelated, especially in the boundary layer. We decided that the best estimate of the effective N and hence standard error of the mean observed mole fraction was the number of vertical profiles contributing to each height bin. We thus constructed vertical profiles of mean observed mole fraction by first averaging all data points in each bin deriving from a common vertical profile, and then weighting all these individual profile averages equally to form the overall bin average mole fraction. The differences in vertical profiles of observed mole fraction found by this method instead of simply averaging all five-second data points within each bin were minor.

For each of the global inversions, vertical profiles of mean modeled mole fraction were found using the same method as for the observations. For a given region, sector, and altitude bin, the relative magnitude of the bias (defined as model minus observation) to the standard error of the observations gives an indication of the significance of that model's bias. It is likely this is a conservative estimate of significance, because much of the variance of the observations (i.e., from profile to profile with our method) is correlated with variance in the modeled profile (i.e., the models have nonzero skill at predicting synoptic variability).

Providing an overall assessment of model biases relative to the observations is problematic because the relatively small sample size and large variability in model properties preclude even a rough application of Gaussian statistical methods. However, to give some indication of model-to-model consistency in biases, we show vertical profiles of the 25^{th} quantile of modeled mole fraction, the 50^{th} quantile (i.e., the median), and the 75^{th} quantile, along with the observed mole fraction for comparison. We also show the ordered rank of the observation as a function of bin, where a rank of 1 indicates the observation has a smaller mole fraction than any of the ten inversions, and a rank of 11 indicates the observation has a larger mole fraction than any inversion.

To help quantify an overall model bias for each of the TM5 / GEOS-Chem inversion groups, in each sector / region combination, we computed median statistics for each transport set of inversions within each 250-m bin as described above. We then averaged the statistics into three layers: below 1500-m AGL, 1500-3000 m AGL, and 3000-4500 m AGL. These layers roughly correspond to the boundary layer (BL) / lower free troposphere (LFT) / upper free troposphere (UFT) classification of Pal et al. (2020) (though in that study observed boundary layer height was used instead of a strict height-based classification). We will henceforth refer to these layers as BL, LFT, and UFT respectively for brevity, with the understanding that these designations are approximate (in particular, BL may or may not correspond to the actual boundary layer).

3 Vertical Profile Summary Statistics

3.1 MidAtlantic (16 Jul -- 31 Jul 2016)

During the Mid-Atlantic phase of the summer campaign, we observed in both warm and cold sectors similar CO_2 vertical profiles above about 1500 m AGL, with CO_2 gradually increasing with height, though the cold sector mole fractions were 2-4 ppm lower on average, and possessed considerably more within-bin and between-bin variability, especially in the UFT (Fig. 1).

Mean inversion posterior CO_2 profiles are broadly consistent with the observations. Above 3000 m AGL, there is little difference between the model median warm and cold sector profiles (compare Fig. 2a and Fig 2b). There is a noticeable overall model negative bias in the upper-level median warm sector compared to observations (up to -2 ppm at 4500 m AGL).

Below 3000 m, the overall bias of median profiles is relatively small (1-2 ppm or less) and the median warm-sector profile captures the observed BL CO_2 enhancement quite well. However, for both sectors the interquantile BL model spread is very large (greater than 5 ppm). Individual inversions can have even greater deviations from the observations (e.g., Fig. 1). The fact that individual model BL variability is far larger than that observed in the UFT strongly suggests that the model spread is due to regional flux variability among the inversions, rather than variability in continental upstream boundary conditions.

While not completely accounting for the overall spread, when the inversions are grouped by transport model, as shown in Fig. 1, systematic transport-related differences, on the order of the model biases, are apparent. Specifically, the TM5 models are lower in CO_2 mole fraction than the GEOS-Chem models. For the warm sector profiles, all of the TM5 models had significant negative CO_2 biases relative to observations below 3000 m AGL (see Fig. 1a), ranging from 2-7 ppm; in contrast, only one other inversion (a GEOS-Chem member) had a significant negative CO_2 bias below 3000 m AGL, and most of the other inversions had significant positive biases. In Table 2, we quantify regional and sector biases as a function of layer and transport model. The corresponding absolute mole fractions are tabulated in Table S2 in the Supplement. We find that the overall TM5 warm sector biases are significantly more negative (-3.9 ppm and -3.1 ppm for the BL and LFT, respectively) than those found in GEOS-Chem (-1.2 ppm and -0.8 ppm). For the cold sectors, the TM5 models are again around 2-5 ppm lower in mole fraction than the GEOS-Chem models below 3000 m AGL, if one outlier GEOS-Chem member is excluded (Fig. 1b); however, in these cases it is the TM5 models which are closer to the observations (BL and LFT biases are -1.3 ppm and -0.6 ppm for TM5, but 2.8 ppm and 2.0 ppm for GEOS-Chem).

It is of interest to examine model skill in reproducing the observed horizontal mole fraction contrast across fronts, as well as the vertical boundary layer / free troposphere mole fraction contrast within each sector, to gain insight into how well the models capture horizontal and vertical transport. To quantify these respective contrasts, we can use the warm sector mole fraction - cold sector mole fraction at each level in Table 2 (henceforth, "sector difference'), and the BL mole fraction - LFT mole fraction for each sector in Table S2 (henceforth "vertical difference"), based on Pal et al. (2020). We also include values of BL mole fraction – UFT mole fraction in Table S2, which can be considered to represent the continental flux signal relative to the continental background. What we find is that in the BL and LFT the observed sector difference (+4.1 ppm and +2.4 ppm) is substantially underestimated by both sets of transport models (TM5: +1.9 ppm and -0.1 ppm; GEOS-Chem: 0.5 ppm and -0.4 ppm). On the other hand, the vertical differences of the observations (-1.4 ppm in warm sectors, -3.5 ppm in cold sectors) are within 1 ppm of each set of transport models (TM5: -2.2 ppm and -4.2 ppm; GEOS-Chem: -1.8 ppm, -2.7 ppm). Thus the models in the MidAtlantic exhibit moderately large, transport-dependent biases in mole fraction and mole fraction sector contrast, relatively consistent up to 3000 m. But the BL – UFT differences are considerably larger for the TM5 models than either GEOS-Chem or the observations (-6.5 ppm warm sectors and -7.2 cold sectors, vs. GEOS-Chem: -4.9 ppm warm and -5.0 ppm cold sectors, observations: -4.3 ppm warm sectors and -5.3 ppm cold sectors). This reflects large depletions of TM5 LFT CO_2 relative to its upper level background.

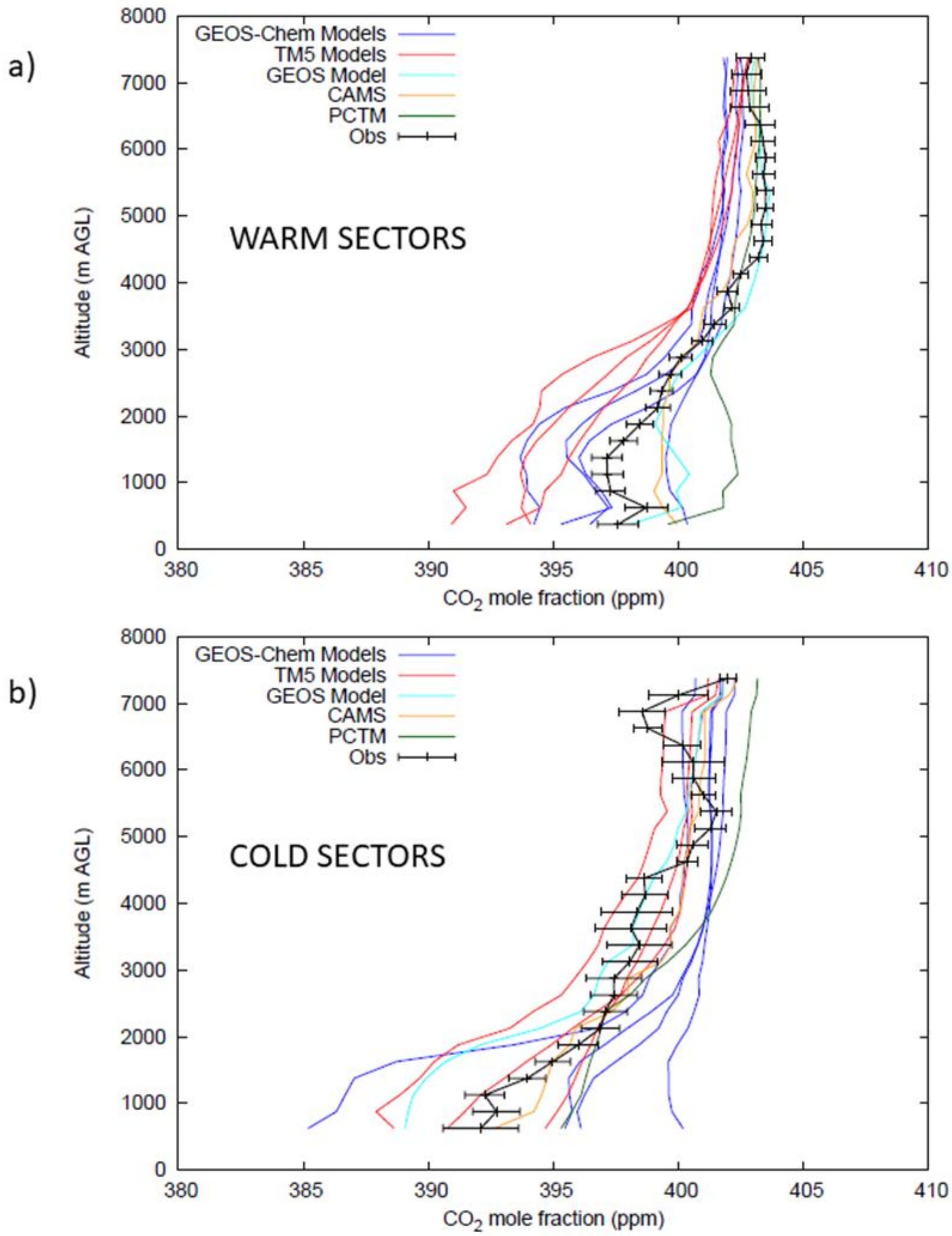

Figure 1. Vertical profiles of mean observed (black) and modeled (colored) mole fraction as a function of altitude AGL for Mid-Atlantic warm sectors (a) and cold sectors (b), from 16-31 Jul 2016. Blue indicates OCO-2 MIP models using GEOS-Chem offline transport; red indicate models using TM5 offline transport; light blue indicates the GEOS-5 inline transport member; orange indicates the CAMS member; green indicates the PCTM transport member. Bars on the

observations indicate $\pm\sigma/\sqrt{N}$, as described in text. Statistics generated from 36 total warm sector profiles and 12 total cold sector profiles.

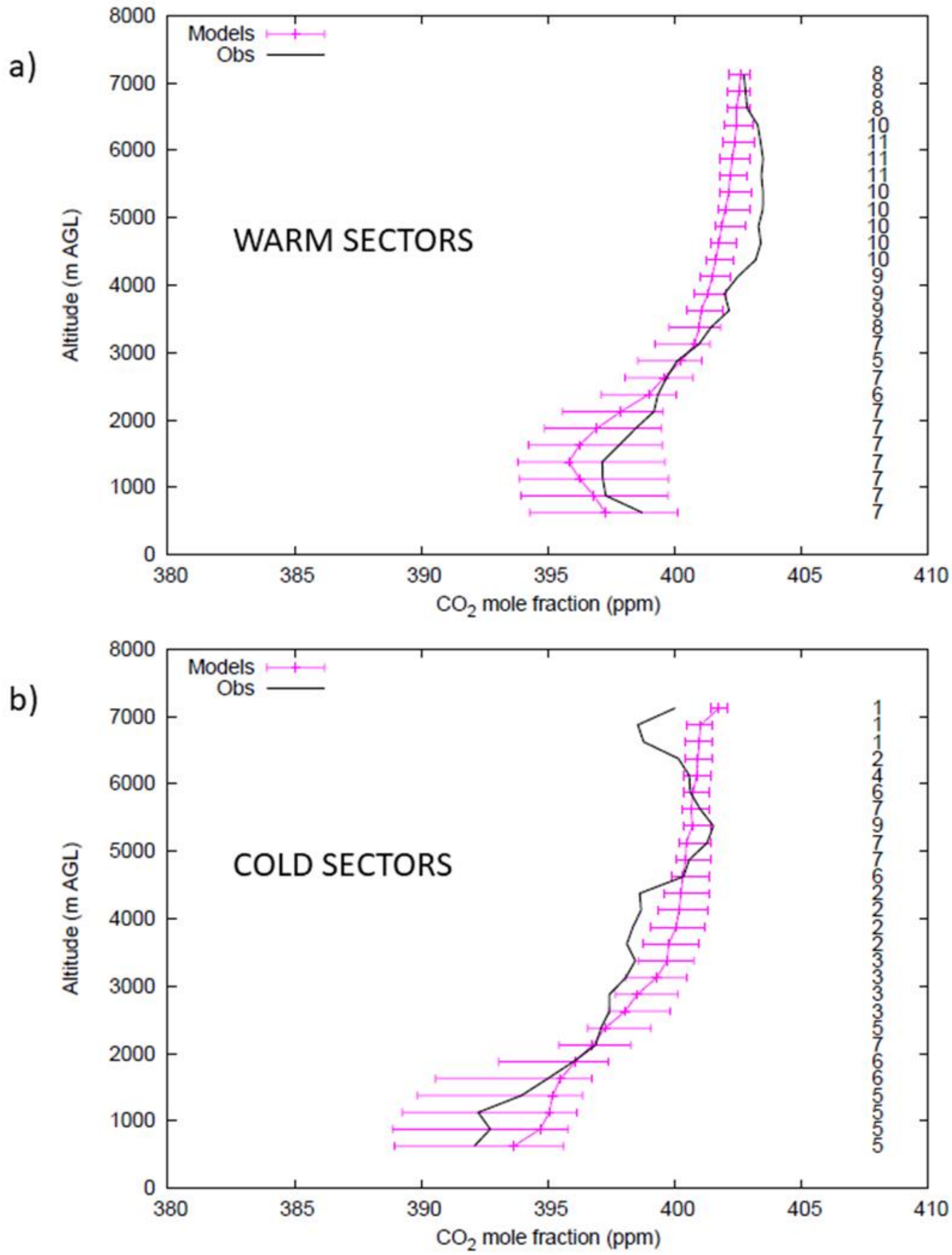

Figure 2. Vertical profiles of mean observed mole fraction (black) and model spread of mole fraction (colored) as a function of altitude AGL for Mid-Atlantic warm sectors (a) and cold sectors (b), from 16-31 Jul 2016. Vertical colored curve indicates median (50th quantile) value of model mole fraction; bars indicate location of 25th and 75th quantile of model mole fraction. Numbers to right indicate ordered rank of observations relative to models, with 1 indicating observed value is less than all 10 model values, and 11 indicating observed value is greater than all 10 model values.

Table 2. Mole fractions biases of CO_2 in layers below 1500 m AGL (BL), 1500 – 3000 m AGL (LFT), and 3000 m – 4500 m AGL (UFT) for both TM5 and GEOS-Chem inversion, within each of the three regions in warm and cold sectors. Also shown are sector differences (warm sector mole fractions minus cold sector mole fractions) for each of GEOS-Chem, TM5, and the observations for each layer and region. Values are the average of each 250-m bin median value within each layer.

Layer	Median TM5 Warm Sector Bias	Median GEOS-Chem Warm Sector Bias	Median TM5 Cold Sector Bias	Median GEOS-Chem Cold Sector Bias	TM5 Sector Difference	GEOS-Chem Sector Difference	Observed Sector Difference
Mid-Atlantic							
Below 1500 m (BL)	-3.9	-1.2	-1.3	+2.8	+1.9	+0.5	+4.5
1500 m – 3000 m (LFT)	-3.1	-0.8	-0.6	+2.0	-0.1	-0.4	+2.4
3000 m – 4500 m (UFT)	-1.7	-0.6	+0.6	+2.5	+1.2	+0.4	+3.5
Midwest							
Below 1500 m (BL)	-0.8	-1.5	+3.6	+3.3	+8.7	+8.3	+13.1
1500 m – 3000 m (LFT)	-0.3	-0.5	-1.6	-2.3	+5.7	+6.2	+4.4
3000 m – 4500 m (UFT)	-0.5	-0.9	-0.6	-0.8	+1.9	+1.7	+1.8
South							
Below 1500 m (BL)	-0.8	-1.3	+2.3	+1.8	+4.0	+4.0	+7.1

1500 m – 3000 m (LFT)	+0.4	0.0	0.0	+0.1	+3.7	+3.2	+3.3
3000 m – 4500 m (UFT)	+0.2	-0.1	-0.1	+0.4	+2.7	+1.9	+2.4

3.2 Midwest (1 Aug -- 16 Aug 2016)

Unlike the MidAtlantic, in the Midwest the observed warm and cold sector profiles show little resemblance to each other. In the warm sectors, (Fig. 3a), little vertical variation of mean CO_2 (1 ppm or less) is seen above 1500 m AGL. Below 1500 m AGL, many (though not all) warm sectors show a few ppm enhancement of CO_2. In contrast, the Midwest cold sectors (Fig. 3b) show the greatest low-level CO_2 depletion relative to upper levels of any observed profile subset (almost 20 ppm), and vertical gradients of the mean CO_2 profile exist throughout the whole atmosphere above 1000 m AGL.

Individual model biases for the Midwest warm sectors are on the order of 3 ppm or less at all levels in the profile, and the bias for the overall median model profile is on the order of 1 ppm. The biases that are present seem to be significant, in that the model spread is even smaller, at least above 1500 m. Between 3500 m and 7000 m AGL, the model mole fractions are too low in virtually every case (Fig. 4a). Below 1000 m AGL the biases are small compared to observational variability (Fig. 3a), but virtually all inversion members have lower mole fractions than the median observed mole fraction, as seen by the high rank of the observations in Fig. 4a. While BL CO_2 enhancements were not observed in every warm-sector case, when they were observed the models underestimated them. The transport model differences appear to be about half of the overall bias in both the UFT and BL (Table 2). But in terms of overall magnitude (< 1 ppm), these transport model differences are substantially less than those in the MidAtlantic.

Though the observational standard errors are large, in the Midwest cold sectors above 1500 m AGL the modeled CO_2 is lower than the observations for virtually all inversions (Fig. 4b). Below 1500 m, the model biases reverse sign, and become positive for virtually all members between 750 – 1250 m AGL. Thus the models systematically underestimate the mole fraction difference between about 1000 and 2500 m AGL. By 3000 m AGL, the Midwest cold sector negative observation-relative biases are reduced in magnitude, and remain relatively constant with additional height.

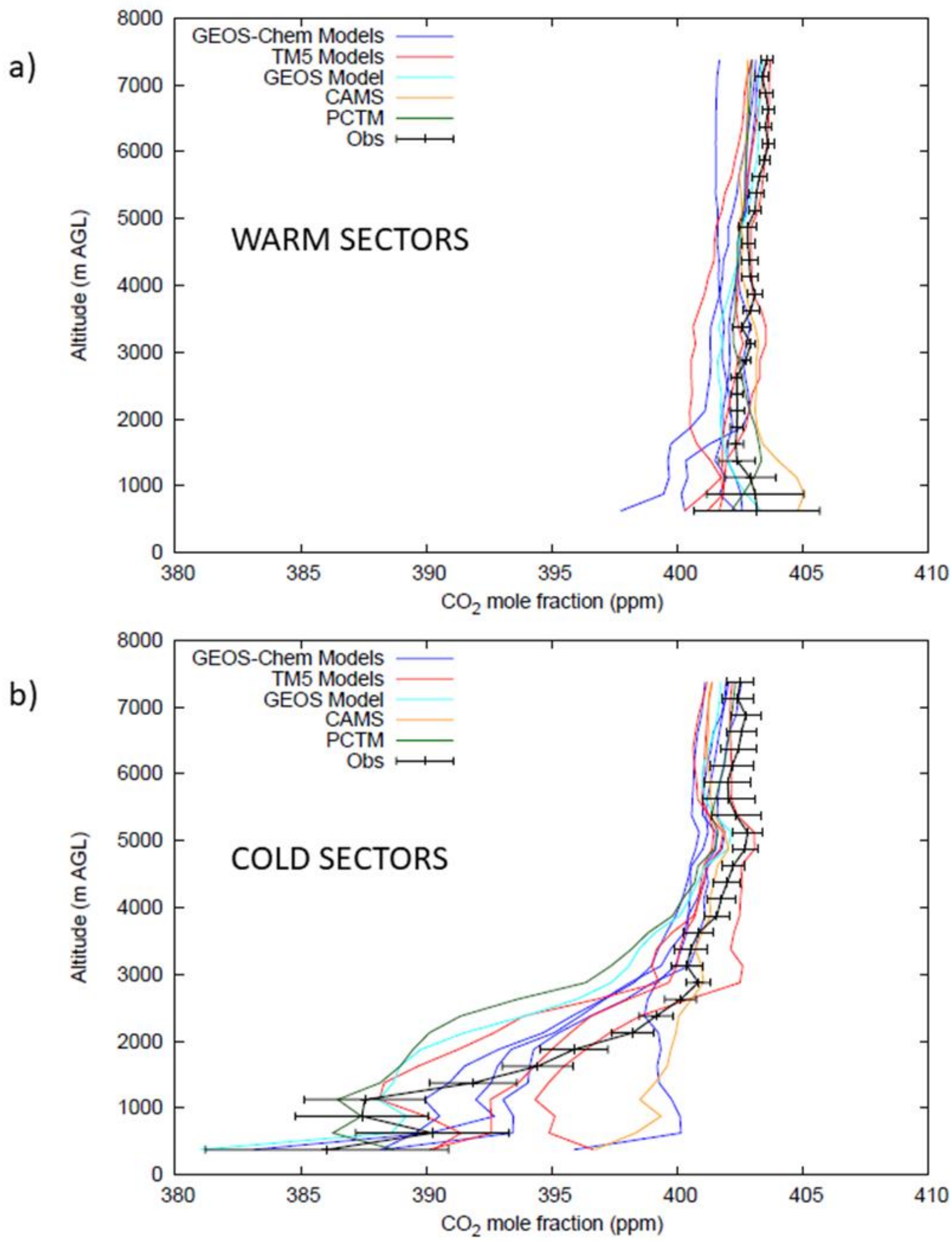

Figure 3: Same as Figure 1, but for the Midwest region (1-15 Aug 2016). Statistics generated from 29 total warm sector profiles and 29 total cold sector profiles.

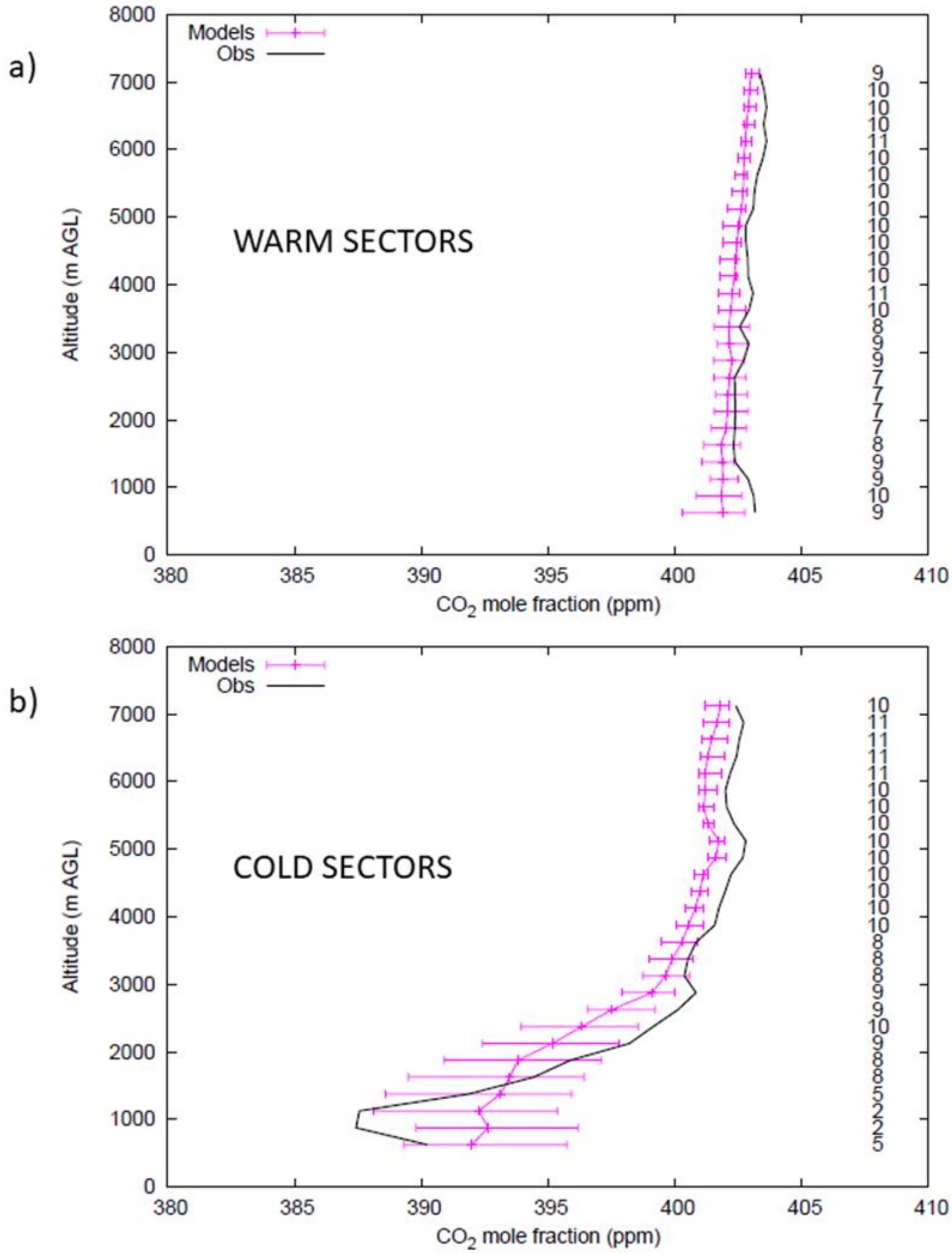

Figure 4: Same as Figure 2, but for the Midwest region (1-15 Aug 2016).

In the BL, the observed Midwest sector difference is much larger in the Midwest than in the Mid-Atlantic (13 ppm). The models underestimate it by about 4 ppm (8.7 ppm in TM5, 9.0 ppm in GEOS-Chem). Interestingly, the Midwest sector difference in the LFT is actually overestimated by the models (5.7 ppm in TM5, 6.2 ppm in GEOS-Chem, vs. 4.4 ppm in the observations). This is most easily explained by the models transporting too much of the cold sector BL depletion signal into the LFT.

These characteristics resemble the growing-season average vertical profiles of the inversions used in Stephens et al. (2007), in which model underestimation of the vertical gradient between 1000 m and 4000 m was used to infer model overestimation of vertical mixing. A similar inference can be made here using BL – LFT (-8.8 ppm observations vs. -3.6 ppm TM5 and -3.2 ppm GEOS-Chem). The reduced observation-relative biases in the UFT vs. the LFT suggest that the LFT communicates more via vertical mixing with the BL than with the UFT. We hypothesize that the enhanced vertical mixing is mainly a result overestimated boundary layer mixing in the models, which would not be expected to exceed 3000 m AGL very frequently.

While the bias of the overall vertical differences relative to the observations are quite large, there is little transport model dependence to this bias. The systematic differences between TM5 and GEOS-Chem transport members in the Midwest cold sectors are even less than they are in the Midwest warm sectors, and are much smaller than the corresponding observational-relative biases. From Table 2, the biases in the BL are -3.6 ppm for TM5 and -3.3 for ppm for GEOS-Chem.

Stephens et al. (2007) went on to show that larger vertical gradients for their seasonal inversion profiles were correlated with decreased magnitudes of net surface posterior fluxes. But that doesn't seem to be the case here. Among the individual inversions, BL mole fractions do not seem to be strongly constrained. The variation of the vertical difference magnitude is large, ranging approximately that of the observations to near zero, and sometimes of opposite sign to the observations – a feature that would be hard to explain solely by vertical mixing differences. Furthermore, while one might expect vertical mixing to be mainly a function of transport model, little dependence of BL – LFT mole fractions on transport model is apparent.

3.3 South (16 Aug - 31 Aug 2016)

In the South warm sectors, there is even less vertical structure in the observed mole fraction profiles above 1500 m AGL than in the Midwest warm sectors (Fig. 5a). Below 1500 m AGL, the clearly observed tendency is for CO_2 mole fractions to be greater than in the free troposphere, by 2-3 ppm. This low-level warm sector enhancement was noted and discussed by Pal et al. (2020), and suggests that a net upwind source of CO_2.

The inversion posterior CO_2 mole fraction profiles for the South warm sectors also show very little vertical variability with height in the mid- and upper-troposphere, with biases that are extremely small in magnitude (< 2 ppm even for the worst-case model above 2 km AGL, and for the median profile on the order of tenths of a ppm at most). But there does seem to be a tendency in Table 2 for the TM5 models to be about 0.3 ppm greater in upper level mole fraction than the GEOS-Chem models, which while small is actually greater than the overall model bias. Nearer the surface, many of the models reproduce but underestimate the enhancement of CO_2 (Fig. 6a).

The number of cold sector profiles in the South was the least numerous of any region, and so the composite profile is somewhat noisy and difficult to interpret (Fig. 5b). Both observed variability and model spread are very small (1 ppm or less) above 3000 m AGL, but extremely large (4-5 ppm) below 3000 m AGL (Fig. 6b). While the median model mole fraction below 1500 m AGL is about 2 ppm greater than the observations, the observed profile lies well within the total model spread (Fig. 6b), and the TM5 and GEOS-Chem biases are within 0.5 ppm of each other (Table 2). The spread of the GEOS-Chem models is substantially larger than for the TM5 models (see Fig. 5b).

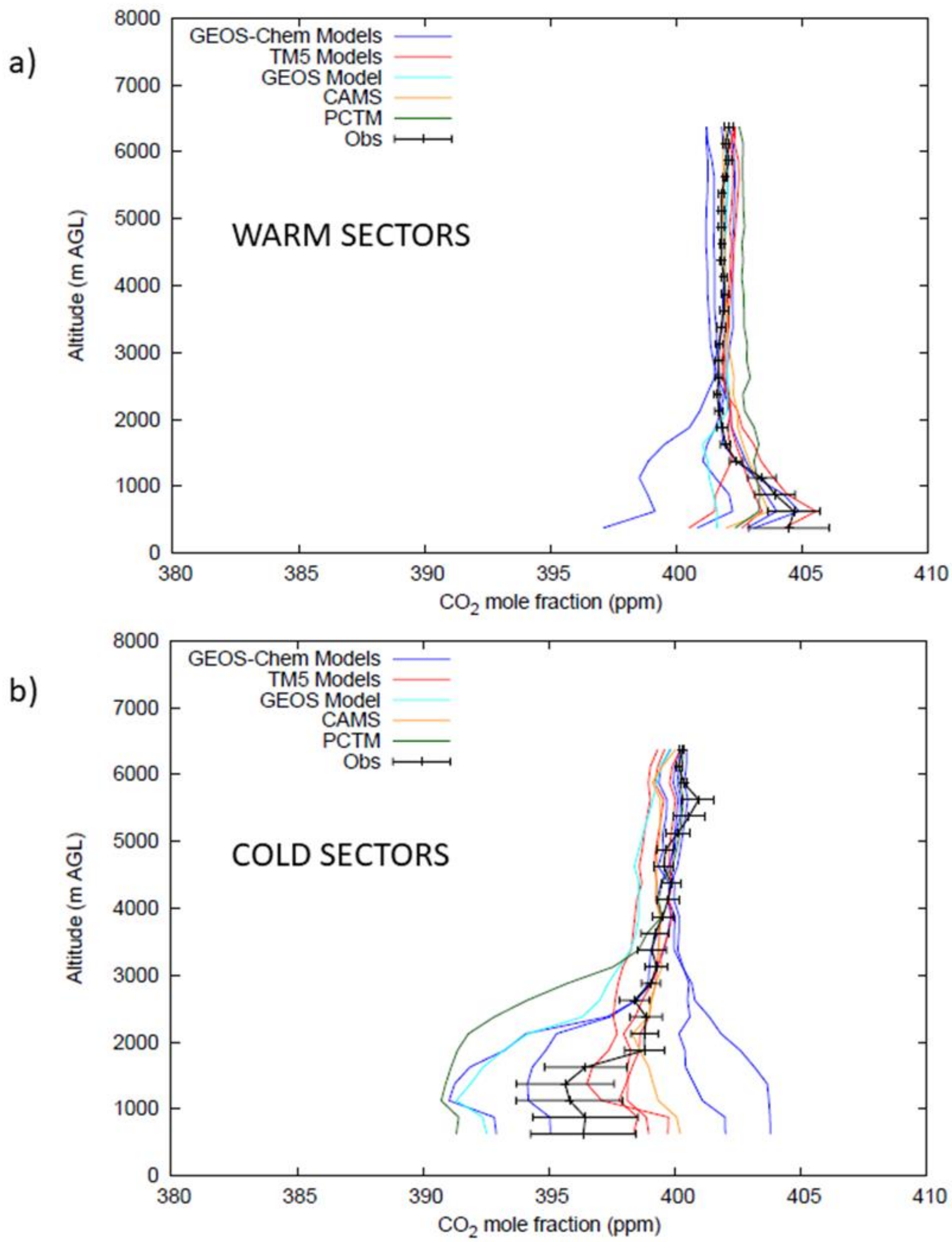

Figure 5: Same as Figure 1, but for the South region (16-31 Aug 2016). Statistics generated from 31 total warm sector profiles and 11 total cold sector profiles.

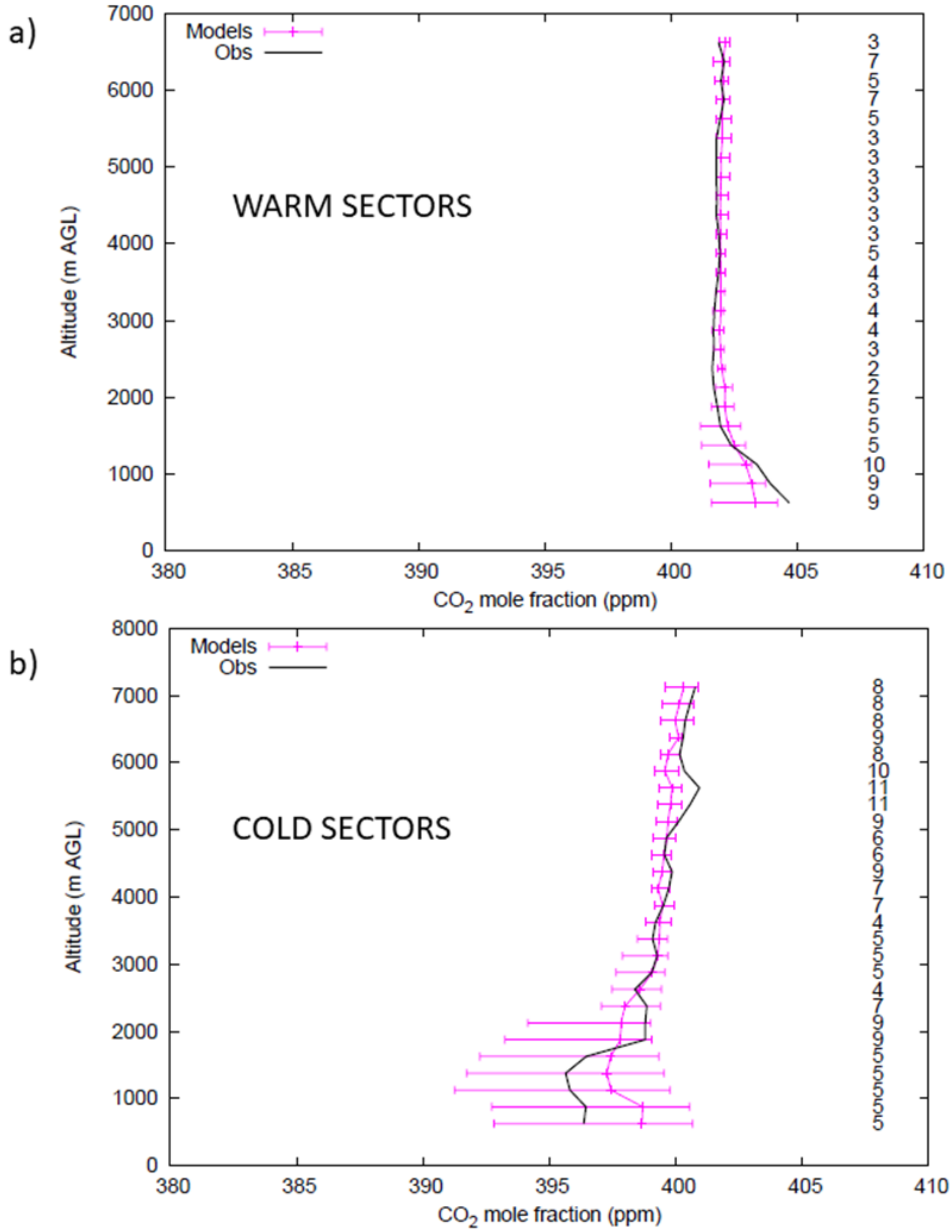

Figure 6: Same as Figure 2, but for the South region (15-31 Aug 2016).

The observed BL sector difference in South (Table 2, 7.1 ppm) is underestimated by both the TM5 and GEOS-Chem models (4 ppm). As in the Midwest, both sets of transport models are biased negative in the BL warm sector by about 1 ppm (-0.8 ppm in TM5, -1.3 ppm in GEOS-Chem), so the larger contribution to the underestimated BL sector gradient is the positive model bias in the cold sector (2.3 ppm in TM5, 1.8 ppm in GEOS-Chem). The magnitude of the observed warm sector vertical difference (+2.0 ppm) is underestimated in both model means (+0.8 ppm for the TM5 models and +0.7 ppm in GEOS-Chem). As in other regions, individual inversion profiles can be very different from the mean, or median, profiles, with some warm sectors even showing low-level depletion.

In the South cold sectors, as was the case in the Midwest, the variation of BL mole fractions and vertical gradients among individual inversions is very large (Fig. 5b), and the median interquantile spread at low levels is on the order of 7 ppm (Fig. 6b). Like the Midwest, the overall median inversion bias for cold sector BL is positive (about +2.0 ppm), although considerably less than the spread among inversions; also like the Midwest, the dependence of this bias on transport model is considerable less than the overall bias itself (i.e., BL cold sector biases of +2.3 ppm for TM5, +1.8 ppm for GEOS-Chem in Table 2). Meanwhile, for the cold sector LFT, inversion biases are virtually non-existent (0.0 ppm for TM5; +0.1 ppm for GEOS-Chem). Though overall the inversion cold sector LFT is unbiased, the observed profile in Fig. 5b shows a sharp drop in mole fraction as one descends from 2000 m to 1500 m AGL that is not present in any of the inversions. As in the Midwest, this suggests that the inversions are overestimating vertical mixing at least to the 2000 m AGL level.

4 Discussion

As documented in Pal et al. (2020), the ACT observations from summer 2016 show BL CO_2 mole fractions are enhanced in warm sectors and depleted in cold sectors throughout all regions. We note that our values for these enhancements / depletions differ from, and in the warm sector case are substantially less than, the values in Pal et al. (2020). This is in part because they included horizontal transects in their analysis, and in part because in their vertical analysis they used boundary layer height rather than 1500 m AGL to define their lowest layer.

Except for the warm sectors of the Midwest and South, the spread among individual inversions is quite large compared to the median inversion biases, especially near the surface. (For Midwest cold sectors, the near-surface model spread is large, but the positive model bias is even larger.) The low-level observational variability is also much less than the total inversion model spread. Thus, in these cases ACT-America observations can serve to select out particular inversion members that have more realistic profiles than the others.

Errors in modeled regional CO_2 mole fraction can be attributed to errors in regional surface fluxes, errors in transport (vertical and horizontal), and/or errors in the regional upstream boundary conditions (e.g., those of the inflow marine air masses to the North American continent). The UFT mole fractions may be used as a proxy for the continental upstream boundary conditions, under the assumption that those boundary conditions are vertically well-mixed, and that the UFT has been little affected by the regional flux signal. These assumptions are consistent with the aircraft climatology from Sweeney et al., (2015) that shows vertically homogeneous summer CO_2 profiles at the west and Gulf coasts, the boundary conditions for most of the air we are sampling, and relatively small mole fraction differences (a few ppm in the BL) between the Gulf and west coast air masses. Our measurements show the Gulf background at least is 401-402 ppm.

Large inversion-to-inversion differences in the posterior CO_2 profiles, with little correlation to atmospheric transport model, suggest that the differences of the regional posterior fluxes among the inversions must be quite large. For cases where the profile model spread near the surface is large (i.e., over 5 ppm), the corresponding UFT model spread is far less (no more than 2 ppm). This (2 ppm) is also comparable in magnitude to UFT bias of the median model relative to the observations. It is clear that the magnitude of the variability among the inversion profiles is too large to be due to variability in their upstream boundary conditions. This variability is also too large to be attributed solely to transport model differences (TM5 / GEOS-Chem differences are 3-4 ppm in the MidAtlantic and even less elsewhere). Therefore, only large variation in the regional surface posterior fluxes among the inversions can explain the profile variability.

The systematic tendencies in the Midwest and South suggest an overestimate in vertical mixing that is common for both the TM5 and GEOS-Chem transport models. These two regions are observed to have enhanced BL mole fractions relative to the LFT in warm sectors, and depleted BL mole fractions in model cold sectors; the models reproduce this tendency, but underestimate the magnitude of the BL / LFT vertical difference. The corresponding BL sectoral differences are underestimated relative to the observations, while the LFT sectoral difference is either unbiased, or overestimated. This suggests that model overestimation of vertical mixing (in particular for the cold sectors) likely contributes to the underestimation of the BL sectoral difference. While it is possible that the inverse flux estimates are biased in such a way as to cause underestimation of the BL sectoral differences (e.g., sinks in the cold sector and sources in the warm sector both underestimated), one would then have to explain how the Midwest cold sector LFT could have biases of the opposite sign as the BL.

The systematic error characteristics of the MidAtlantic are quite different from those of the other regions. While in the Mid-Atlantic some enhancement of near-surface CO_2 in warm sectors can be seen in the observations (near 750 m AGL in Fig. 1a), the vertical profiles of warm and cold sectors are more similar to each other than they are in the other regions. For the models, the warm and cold sector profiles are even more similar to each other – in fact, for both TM5 and GEOS-Chem median models in Table 2, LFT warm sectors actually have lower mole fractions than cold sectors. As noted previously, contrary to other regions the GEOS-Chem – TM5 differences in the MidAtlantic are consistently positive and significant (2-4 ppm in the BL and LFT; 1-2 ppm in the UFT) and exceed the median model sectoral differences, Consequently, the TM5 models have large negative mole fraction biases at all levels in warm sectors, while the GEOS-Chem models have large positive mole fraction biases at all levels in cold sectors.

Also unlike the other regions, in the MidAtlantic the model BL - LFT vertical differences are relatively unbiased, even when the mole fractions in the BL and LFT separately are biased. This does not necessarily mean that the model vertical mixing is correct. It does suggest that the upstream source of the model biases seen in the MidAtlantic has had time to vertically mix up to at least the LFT. This implies that the source of the error is either in the upstream boundary conditions, or in regional fluxes far enough upstream to have a MidAtlantic LFT signal. North American uptake signals that progressively increase west-to-east have been noted previously (Sweeney et al. 2015; Lan et al. 2017; Chen et al. 2019).

For the MidAtlantic warm sectors, the UFT mole fraction biases are of the same sign as those of the LFT, but reduced in magnitude, especially for the TM5 models. But for GEOS-Chem a large part of the biases could be due to errors in the background. For TM5, the value of LFT – UFT is much more negative than that of either GEOS-Chem or the observations (-4.3 ppm vs. -3.1 ppm GEOS-Chem, -2.9 ppm observations), which leads to TM5 having BL – UFT much more negative as well (-6.5 ppm vs. -4.9 ppm GEOS-Chem, -4.3 ppm observations). This may seem to contradict the results in Schuh et al. (2019) that suggested TM5 had more vertical mixing than GEOS-Chem. However, if little local vertical mixing happens between the LFT and UFT in either models or observations, then increased vertical gradients between them imply little further about local vertical mixing. Rather, they would suggest in the MidAtlantic warm sectors that TM5 has too much upstream vertical mixing between the LFT and BL near the region of upstream surface uptake.

For the MidAtlantic cold sectors, the positive GEOS-Chem biases are again quite consistent with height up to the UFT, suggesting most of the (large) biases are in the background. TM5 has reduced mole fraction biases, but again has more negative LFT - UFT and BL – UFT than GEOS-Chem or the aircraft (LFT - UFT: -3.0 ppm for TM5 vs. -2.3 ppm for GEOS-Chem and -1.8 ppm for observations; BL – UFT: -7.2 ppm vs. -5.0 ppm and -4.9 ppm). We would again argue that the TM5 vs. GEOS-Chem vertical gradient differences reflect more upstream vertical mixing of a surface uptake signal to the LFT in TM5. The observational vertical gradient might suggest the vertical mixing of GEOS-Chem is more realistic, but with much large positive background bias. But it should also be noted that the observational variability of MidAtlantic cold sectors in the UFT and higher is large, with different layers ranging from 398 – 401 ppm. If the ‘true’ upper level background were 401 ppm, one could well conclude that TM5 has the more realistic vertical mixing to go along with a large negative background bias.

To better understand the differences between regions, we also did backward trajectories using the NOAA Hybrid Single Particle Lagrangian Integrated Trajectory (HYSPLIT) model (Stein et al., 2015; Rolph et al., 2017) from the location of soundings from each combination of sector and region. This is similar to the procedure of Pal et al. (2020), except that we took the trajectories back to 120 hours, and for each sounding location we start backward trajectories at 1000 m, 2000 m, and 4000 m AGL, to represent the BL, LFT, and UFT layers. To minimize the sampling uncertainty associated with initiating a single backwards trajectory, at each sounding location we initiate a 3 x 3 matrix of backwards trajectories for a 0.2 x 0.2 degree square centered at the sounding location. In the Supplement, we show representative trajectories from each region / sector combination.

Though considerable variability exists among back-trajectories within each region, in general for the Midwest and South warm sectors they originate from the Gulf of Mexico at all levels, as might be expected from their vertical homogeneity. For the cold sectors of these regions, the back-trajectories tend to derive from the north or northwest at low levels, and from southwesterly to northwesterly directions at higher levels. For the MidAtlantic, warm sector back-trajectories, we found parcels from about 1000 m AGL tend to derive from the U.S. Southeast and Gulf Coast, or the Atlantic; at higher levels, parcels derive from more westerly directions, the

Appalachians and upper Midwest (Ohio and Mississippi River Valleys). The MidAtlantic cold sector back trajectories are generally from the west or northwest at all levels.

The relatively small MidAtlantic airmass sectoral differences (at least above 1000 m AGL) are plausibly a consequence of the relatively similar origins of the back-trajectories for warm sector and cold sector air. The LFT back-trajectories indicate that for both warm and cold sectors the origin is in the eastern Midwest (Mississippi River Valley), a region with evidence of large growing season uptake (Miles et al., 2012; Lokupitiya et al., 2016). So one hypothesis for the biases in the MidAtlantic is that they are driven by either biases in this source region's uptake flux or biases in the transport of this uptake signal. By contrast, warm sector and cold sector air masses in the other regions may be little affected by upstream sources (e.g., they arrive from the Gulf, the Pacific, or the sparsely-vegetated U.S. West), so their BL and LFT fluxes are being largely driven by flux signals from the immediate vicinity.

The observations suggest the TM5 representation of how the MidAtlantic cold sectors become depleted in CO_2 is better which may indicate that the TM5 representation of upstream vertical mixing is more realistic. However, the constancy of the cold sector GEOS-Chem positive bias with height suggests that the GEOS-Chem errors are not so much due to vertical mixing errors, but to not properly representing the general flow from the Midwest surface sinks to the MidAtlantic. Similarly, the large TM5 negative biases for the warm sectors could derive from improper advection of the Midwest depletions in these scenarios, especially near the surface where we found the general flow was from the Gulf Coast states. But it may also be that the shallow nature of many MidAtlantic warm sectors BLs may make them particularly difficult for inversions with too much vertical mixing to represent. More definite assessment of the skill of transport model vertical mixing may require more disentangling of model vertical and horizontal transport errors.

Horizontal transport dependence could be most important in the MidAtlantic because of the relatively small differences between warm and cold sector trajectories, which make it difficult to properly represent low-level air streams in inversions with the resolution of global models (i.e., 1 deg x 1 deg for CT-NRT, and coarser for the other inversions), as well as topography (the Appalachians) not found in other regions. But sharp boundaries between air masses may still exist in these scenarios, such that small horizontal transport errors may lead to substantial mole fraction errors. Experiments using inversions with a common horizontal resolution might help to reduce some of the horizontal transport differences.

Another outstanding issue, if increased TM5 vertical mixing relative to GEOS-Chem is in fact responsible for its larger LFT CO_2 depletion in the MidAtlantic, is why in the Midwest no transport-dependent vertical mixing differences were apparent. Possibly vertical mixing for the source regions of MidAtlantic and Midwest surface uptake just has different characteristics. Or it could be that local boundary-layer mixing in the Midwest is overestimated by all models, but downstream vertical mixing (e.g., by convection) may be transport-model-dependent. Even if boundary layer mixing were in fact transport-model dependent, the fact that the posterior fluxes are not consistent across all transport members could act to suppress clear transport-dependent signals in inversion vertical mixing. A more controlled experiment such as that in Schuh et al. (2019), where only transport model, or surface flux, is varied would lead to more understanding of inversion biases for the ACT-America cases. Current ongoing research, including direct comparison of the simulated vs. observed ACT boundary layer heights, and more controlled modeling experiments such as more rigorous transport model comparison tests with SF_6, will hopefully provide clarification to the realism of vertical mixing in the transport models.

In the warm sectors, BL enhancements of CO_2 were observed in all regions, corresponding to trajectories from the Gulf Coast. (In the MidAtlantic, enhancements were seen in the lowest levels of the warm sectors, possessing Gulf Coast origins; at slightly higher levels, depleted mole fractions were present in warm sectors for trajectories that had a more Midwestern origin.) The models reproduced the enhancement to some degree, suggesting they had local positive net CO_2 fluxes, but generally underestimated it. By contrast, Schuh et al. (2019) reported that these OCO-2 inversions had negative mean net CO_2 fluxes between the equator and 45 N during Aug 2016. However, they also found the fluxes equatorward of 45 N to be substantially reduced from those poleward of 45 N at this time, and the poleward of 45 N seasonal net uptake was rapidly declining (see also Crowell et al., 2019). It is possible that by the time of summer 2016 campaign, the zone of maximum net surface uptake had moved northward from the Gulf Coast states, leading to higher average mole fractions in the northward-moving warm sectors than the southward-moving cold sectors over the ACT focus regions. Ongoing analysis of the last ACT-America campaign from July 2019 should help to address this question.

5 Conclusions

We have shown that the differences in inversion methodology and/or prior fluxes, as opposed to choice of transport model, are the dominant source of variability among the version 7 OCO-2 MIP *in situ* inversion systems analyzed here. In addition, we have shown that the ACT campaign aircraft observations are clearly able to identify biases, and thus are a valuable tool for testing flux inversions. We have found, after segregating models and observations by altitude and airmass within synoptic weather systems, that biases in posterior CO_2 among the individual inversions can be very large compared with the standard error of the observations, suggesting that these biases are significant. Thus, the posterior CO_2 fluxes from these inversions are also likely systematically biased by large amounts.

The median of the OCO-2 MIP4 *in situ* inversions, while containing some systematic biases, captures the observed CO_2 profile characteristics in different regions and sectors fairly well. This analysis suggests the advantage of using central estimates from a suite of models such as the OCO-2 MIP, as opposed to any single member.

The significant systematic biases in the model median mean cross-frontal and LFT-ABL posterior CO_2 differences in the Midwest and South suggest some caution in using the central estimates. These biases are likely a combination of underestimate of cold sector net uptake, warm sector net source, and an overestimate of BL-LFT vertical mixing in the U.S. midcontinent. These biases do not appear to be related to the differences in the TM5 vs. GEOS-Chem atmospheric transport.

Significant differences between TM5 vs. GEOS-Chem based inversions in the MidAtlantic point toward the importance of deeper tropospheric mixing farther downwind in the continent. They suggest that TM5 has greater vertical mixing of the upstream depletion signal reflected in much of the MidAtlantic cold sector profile, and may suggest that TM5 models this process more realistically. These conclusions are similar to Schuh et al., (2019), who found more limited vertical mixing vs. horizontal mixing in GEOS-Chem relative to TM5. However, TM5 model errors are much larger than those of GEOS-Chem for MidAtlantic warm sectors, possibly due to horizontal transport errors or excessive local TM5 mixing near the surface in these situations. Biases in inversions based on both atmospheric transport models make it difficult to reach a definitive conclusion.

Because members of the version 7 OCO-2 MIP suite consist of different permutations of prior fluxes and transport models, clear attribution of these systematic differences is not possible. Experiments that can more fully diagnose the origin of the posterior CO_2 biases and provide insight into which inversions provide more realistic posterior CO_2 flux estimates are needed. This can be done with controlled experiments using common surface fluxes but different transport models (e.g. Diaz-Isaac et al, 2014; 2018; Chen et al., 2019; Schuh et al., 2019), and common atmospheric transport but different fluxes (e.g. Feng et al., 2019). Further investigation about the cause of divergence among fluxes can be achieved with comparisons of aircraft CO_2 to inversions that use the same atmospheric CO_2 observations but different inversion methods, or the same inversion methods but different observations – including of course the use of OCO-2 XCO_2 retrievals, a key component of the OCO-2 MIP itself (Crowell et al., 2019). Efforts to expand ACT model-data comparisons in all of these directions, and to include evaluation of atmospheric winds and BL depth (e.g. Diaz-Isaac et al., 2018), are underway. Simulations (Feng et al., 2019) and observational methods (Baier et al., 2019) that can differentiate the origins of continental CO_2 provide yet another tool for diagnosing model performance. Finally, comparisons that span all four seasons will be needed to gain insight into annual flux inversion results. These experiments are all underway as part of the analysis of the ACT-America flight campaigns.

Acknowledgments, Samples, and Data

This work was sponsored by the National Aeronautics and Space Administration (NASA) under award NNX15AG76G. ACT-America is a NASA Earth Venture Suborbital – 2 mission supported by NASA's Earth Sciences Directorate. The data used are freely available from the NASA archive (https://www-air.larc.nasa.gov/cgi-bin/ArcView/actamerica.2016) and the ORNL Distributed Active Archive Center (https://doi.org/10.3334/ORNLDAAC/1574), and the model output along the flight tracks is available in Obspack format (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/OCO2_insitu/). This research would not have been possible without the hard work of a great number of people from multiple institution; we would especially like to thank NASA headquarters and the Airborne Sciences program for the logistics of site deployment, and the pilots for skillfully following the flight plans. The authors gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model and READY website (http://www.ready.noaa.gov) used to generate the back-trajectories in the Supplemental Information. Co-author SP was supported by NASA Grant Number 80NSSC19K0730 and a Texas Tech University start up research grant. All the authors declare no financial conflict of interests. We thank two anonymous reviewers for their constructive comments.

References

Agustí-Panareda, A., Massart, S., Chevallier, F., Balsamo, G., Boussetta, S., Dutra, E., & Beljaars, A. (2016, August). A biogenic CO_2 flux adjustment scheme for the mitigation of large-scale biases in global atmospheric CO_2 analyses and forecasts. *Atmospheric Chemistry and Physics*, *16*(16), 10399–10418. doi: 10.5194/acp-16-10399-2016.

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., et al. (2014). CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts. *Atmospheric Measurement Techniques*, *7*(2), 647–687. https://doi.org/10.5194/amt-7-647-2014

Baier, B.C., C. Sweeney, Y. Choi, K.J. Davis, J.P. DiGangi, S. Feng, A. Fried, et al., 2019: Multispecies assessment of factors influencing regional CO_2 and CH_4 enhancements during the winter 2017 ACT-America campaign. *J. Geophys. Res. Atmospheres*, **125**, e2019JD031339. https://doi.org/10.1029/2019JD031339.

Baker, D., R. Law, K. Gurney, P. Rayner, P. Peylin, A. Denning, P. Bousquet, L. Bruhwiler, Y. Chen, P. Ciais, I. Fung, M. Heimann, J. John, T. Maki, S. Maksyutov, K. Masarie, M. Prather, B. Pak, S. Taguchi, and Z. Zhu, 2006: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988-2003. *Global Biogeochem. Cycles,* **20**, GB1002, doi:10.1029/2004GB002439.

Baker, D.F., H. Bösch, S.C. Doney, D. O'Brien, and D.S. Schimel, 2010: Carbon source/sink information provided by column CO_2 measurements from the Orbiting Carbon Observatory. *Atmos. Chem. Phys*., **10**, 4145-4165. doi:10.5194/acp-10-4145-2010.

Ballantyne, A.P., C.B. Alden, J.B. Miller, P.P. Tans, and J.W.C. White, 2012: Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*, **488**, 70-72.

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

Basu, S., D.F. Baker, F. Chevallier, P.K. Patra, J. Liu, and J.B. Miller, 2018: The impact of transport model differences on CO_2 surface flux estimates from OCO-2 retrievals of column average CO_2. *Atmos. Chem. Phys*., **18**, 7189.

Berrisford, P., D. Dee, P. Poli, R. Brugge, K. Fielding, M. Fuentes, P. Kallberg, S. Kobayashi, S. Uppala, and A. Simmons, 2011: The ERA-Interim archive; Version 2.0. Retrieved from https://www.ecmwf.int/en/elibrary/8174-era-interim-archive-version-20.

Bey, I., D.J. Jacob, R.M. Yantosca, J.A. Logan, B.D. Field, A.M. Fiore, Q.-B. Li, H. Liu, L.J. Mickley, and M.G. Schultz, 2001: Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *J. Geophys. Res*., **106(D19)**, 23073-23095.

Bosilovich, M.G., S. Akella, L. Coy, R. Cullather, C. Draper, R. Gelaro, R. Kovach, Q. Liu, A. Mold, P. Norris, and coauthors, 2015: Technical report series on global modeling and data assimilation, volume 43. MERRA-2: Initial evaluation of the climate, NASA-GMAO. Retrieved from https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf.

Chen, H.W., L.N. Zhang, F. Zhang, K.J. Davis, T. Lauvaux, S. Pal, B. Gaudet, and J.P. DiGangi, 2019: Evaluation of regional CO_2 mole fractions in the ECMWF CAMS real-time atmospheric analysis and NOAA CarbonTRacker Near-REal Time reanalysis with airborne observations from ACT-America field campaigns. *J. Geophys. Res. Atmospheres*, **124,** 8119-8133.

Chevalier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F.-M. Bréon, A. Chédin, and P. Ciais, 2005: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *J. Geophys. Res*., **110**, D24309, doi:10.1029/2005JD006390.

Chevalier, F., L. Feng, H. Bösch, P.I. Palmer, and P.J. Rayner, 2010: On the impact of transport model errors for the estimation of CO_2 surface fluxes from GOSAT observations. *Geophys. Res. Lett*., **37**, L21803-L21803, http://dx.doi.org/10.1029/2010GL044652DO.

Crisp, D., 2015: Measuring atmospheric carbon dioxide from space with the Orbiting Carbon Observatory-2 (OCO-2). *Proc. SPIE,* 9607, https://doi.org/10.1117/12.2187291.

Crowell, S., S.R. Kawa, E. Browell, D. Hammerling, B. Moore, K. Schaefer, and S. Doney, 2018: On the ability of space-based passive and active remote sensing observations of CO_2 to detect flux perturbations to the carbon cycle. *J. Geophys. Res. Atmospheres*, doi:10.1002/2017JD027836.

Crowell, S., D. Baker, A. Schuh, S. Basu, A.R. Jacobson, F. Chevallier, J. Liu, F. Deng, L. Feng, A. Chatterjee, D. Crisp, A. Eldering, D.B. Jones, K. McKain, J. Miller, R. Nassar, T.

Oda, C. O'Dell, P.I. Palmer, D. Schimel, B. Stephens, and C. Sweeney, 2019: The 2015-2016 carbon cycle as seen from OCO-2 and the global *in situ* network. Atmospheric Chemistry and Physics Discussions. https://doi.org/10.5194/acp-2019-87

Davis, K.J., M.D. Obland, B. Lin, T. Lauvaux, C. O'Dell, B. Meadows, E.V. Browell, J.P. DiGangi, C. Sweeney, M.J. McGill, J.D. Barrick, A.R. Nehrir, M.M. Yang, J.R. Bennett, B.C. Baier, A. Roiger, S. Pal, T. Gerken, A. Fried, S. Feng, R. Shrestha, M.A. Shook, G. Chen, L.J. Campbell, Z.R. Barkley, and R.M. Pauly. 2018. ACT-America: L3 Merged In Situ Atmospheric Trace Gases and Flask Data, Eastern USA. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1593

Deng, F., and J.M. Chen, 2011: Recent global CO_2 flux inferred from atmospheric CO_2 observations and its regional analyses. *Biogeosciences*, **8**, 3263-3281. doi:10.5194/bg-8-3263-2011.

Denning, A.S., I.Y. Fung, and D. Randall, 1995: Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota. *Nature*, **376**, 240-243. doi:10.1038/376240a0.

Díaz Isaac, Liza I., T. Lauvaux, K.J. Davis, N. L. Miles, S. J. Richardson, A. R. Jacobson, and A. E. Andrews, 2014. Model-data comparison of MCI field campaign atmospheric CO_2 mole fractions. *Journal of Geophysical Research – Atmospheres*, 119, 10,536-10,551, doi:10.1002/2014JD021593.

Díaz-Isaac , Liza I., T. Lauvaux, K.J. Davis: Impact of physical parameterizations and initial conditions on simulated atmospheric transport and CO_2 mole fractions in the US Midwest. *Atmos. Chem. Phys*., 18, 14813–14835, doi.org/10.5194/acp-18-14813-2018, 2018.

Eldering, A., C. W. O'Dell, P.O. Wennberg, D. Crisp, M.R. Gunson, C. Viatte, C. Avis, A. Braverman, R. Castano, A. Chang, and coauthors, 2017: The orbiting carbon observatory-2: First 18 months of science data products. *Atmospheric Measurement Techniques*, **10(2)**, 549–563.

Feng, L., P.I. Palmer, N.M. Deutscher, D.G. Feist, I. Morino, and R. Sussmann, 2016: Estimates of European uptake of CO_2 inferred from GOSAT XCO2 retrievals: sensitivity to measurement bias inside and outside Europe. *Atmos. Chem. Phys*., **16**, 1289-1302.

Feng, S., T. Lauvaux, K.J. Davis, K. Keller, Y. Zhou, C. Williams, A.E. Schuh, J. Liu, and I. Baker, 2019: Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO_2 simulations over North

America. *J. Geophys. Res. Atmospheres*, **124**, 14325-14346. https://doi.org/10.1029/2019JD031165.

Gaubert, B., B.B. Stephens, S. Basu, F. Chevallier, F. Deng, E.A. Kort, P.K. Patra, W. Peters, C. Rodenbeck, and coauthors, 2019: Global atmospheric CO_2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate. *Biogeosciences*, **16**, 117-134. https://doi.org/10.5194/bg-16-117-2019.

Gurney, K.R., Law, R.M., Denning, A.S., Rayner, P.J., Pak, B.C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I.Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Peylin, P., Prather, M., and Taguchi, S., 2004: Transcom 3 inversion intercomparison: model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochem. Cycles*, **18**, GB1010, doi:10.1029/2003GB002111.

Held, I.M., and T. Schneider, 1999: The surface branch of the zonally averaged mass transport circulation in the troposphere. *J. Atmos. Sci*., **56**, 1688-1697.

Krol, M., S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener, and P. Bergamaschi, 2005: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmos. Chem. Phys*., **5**, 417-432.

Lan, X., P. Tans, C. Sweeney, A. Andrews, A. Jacobson, M. Crotwell, E. Dlugokencky, J. Kofler, P. Lang, K. Thoning, and S. Wolter, 2017: Gradients of column CO_2 across North America from the NOAA Global Greenhouse Gas Reference Network. *Atmos. Chem. Phys*., **17**, 15151-15165, https://doi.org/10.5194/acp-17-15151-2017.

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

Lokupitiya, E., A.S. Denning, K. Schaefer, D. Ricciuto, R. Anderson, M.A. Arain, I. Baker, A.G. Barr, G. Chen, and coauthors, 2016: Carbon and energy fluxes in cropland ecosystems: a model-data comparison. *Biogeochemistry*, doi: 10.1007/s10533-016-0219-3.

Miles, N.L., S.J. Richardson, K.J. Davis, T. Lauvaux, A.E. Andrews, T.O. West, V. Bandaru, and E.R. Crosson, 2012: Large amplitude spatial and temporal gradients in atmospheric boundary layer CO_2 mole fractions detected with a tower-based network in the U.S. upper Midwest. *J. Geophys. Res*., **117**, G01019. doi:10.1029/2011JG001781.

Miles, N.L., S.J. Richardson, D.K. Martins, K.J. Davis, T. Lauvaux, B.J. Haupt, and S.K. Miller, 2018: ACT-America: L2 In Situ CO_2, CO, and CH_4 concentrations from towers, Eastern USA. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1568.

Moore, B., S.M.R. Crowell, P.J. Rayner, J. Kumer, C. W. O'Dell, D. O'Brien, S. Utembe, I. Polonsky, D. Schimel, and J. Lemen, 2018: The potential of the Geostationary Carbon Cycle Observatory (GeoCarb) to provide multi-scale constraints on the carbon cycle in the Americas. *Front. Env. Sci*., **6**, 109, doi:10.3389/fenvs.2018.00109.

Pal, S., 2019: ACT-America: Profile-based planetary boundary layer heights, Eastern USA. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1706.

Pal, S., K.J. Davis, T. Lauvaux, E.V. Browell, B.J. Gaudet, D.R. Stauffer, M.D. Obland, Y. Choi, J.P. Digangi, S. Feng, B. Lin, N.L. Miles, R.M. Pauly, S.J. Richardson, and F. Zhang, 2020: Observations of greenhouse gas changes across summer frontal boundaries in the Eastern United States. *J. Geophys. Res. Atmospheres*, 125, e2019JD030526. https://doi.org/10.1029/2019JD030526.

Pauluis, O., A. Czaja, and R. Korty, 2009: The global atmospheric circulation in moist isentropic coordinates. *J. Climate*, **23**, 3077-3093.

Parazoo, N.C., A.S. Denning, J.A. Berry, A. Wolf, D.A. Randall, S.R. Kawa, O. Pauluis, and S.C. Doney, 2001: Moist synoptic transport of CO_2 along the mid-latitude storm tack. *Geophys. Res. Letters*, **38**(9).

Peixoto, J.P., and Oort, A.H., 1992: *Physics of Climate*, 1st edition. American Institute of Physics, New York, 520 pp.

Peters, W., A. R. Jacobson, C. Sweeney, A. E. Andrews, T.J. Conway, K. Masarie, J.B. Miller, L.M. P. Bruhwiler, G. Petron, A.I. Hirsch, D.E.J. Worthy, G.R. van der Werf, J.T. Randerson, P.O. Wennberg, M.C. Krol, and P.P. Tans, 2007: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 18,925–18,930. https://doi.org/10.1072/pnas.07089861074.

Peylin, P., R.M. Law, K.R. Gurney, F. Chevallier, A.R. Jacobson, T. Maki, Y. Niwa, P.K. Patra, W. Peters, P.J. Rayner, C. Rödenbeck., I.T. van der Laan-Luijkx, and X. Zhang, 2013:

Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions. *Biogeosciences,* **10**, 6699-6720. doi:10.5194/bg-10-6699-2013.

Reinecker, M.M.,, M.J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R.D. Koster, R. Gelaro, I. Stajner, and J.E. Nielsen, 2008: The GEOS-5 Data Assimilation System – Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. Tech. Report, NASA Goddard Space Flight Center.

Rolph, G., A. Stein, A., and B. Stunder, 2017: Real-time Environmental Applications and Display sYstem: READY. Environmental Modelling & Software, **95**, 210-228, https://doi.org/10.1016/j.envsoft.2017.06.025. (http://www.sciencedirect.com/science/article/pii/S1364815217302360)

Schimel, D., Stephens, B.B., and J.B. Fisher, 2015: Effect of increasing CO_2 on the terrestrial carbon cycle. *Proc. National Academy of Sciences*, **112(2)**, 436-441.

Schuh, A., A. Denning, K. Corbin, I. Baker, M. Uliasz, N. Parazoo, A. Andrews, and D. Worthy, 2010: A regional high-resolution carbon flux inversion of North America for 2004. *Biogeosciences*, 7, 1625-1644.

Schuh, A.E., A.R. Jacobson, S. Basu, B. Weir, D. Baker, K. Bowman, F. Chevallier, S. Crowell, K. Davis, F. Deng, S. Denning, L. Feng, D. Jones, J. Liu, and P. Palmer, 2019: Quantifying the impact of atmospheric transport uncertainty on CO_2 surface flux estimates. *Global Biogeochemical Cycles*, https://doi.org/10.1029/2018GB006086.

Stein, A.F., Draxler, R.R, Rolph, G.D., Stunder, B.J.B., Cohen, M.D., and Ngan, F., 2015: NOAA's HYSPLIT atmospheric transport and dispersion modeling system, *Bull. Amer. Meteor. Soc*., **96**, 2059-2077, http://dx.doi.org/10.1175/BAMS-D-14-00110.1

Stephens, B.B., Long, M.C., Keeling, R.F., Kort, E.A., Sweeney, C., Apel, E.C., et al., 2018: The O_2/N_2 Ratio and CO_2 Airborne Southern Ocean study. *Bull. Amer. Meteorol. Soc*., **99**, 381-402.

Stephens, B. B., K.R. Gurney, P.P. Tans, C. Sweeney, W. Peters, L. Bruhwiler, P. Ciais, M. Ramonet, P. Bousquet, T. Nakazawa, S. Aoki, T. Machida, G. Inoue, N. Vinnichencko, J. Lloyd, A. Jordan, M. Heimann, O. Shibistova, R.L. Langenfelds, L.P. Steele, R.J. Francey, and A.S. Denning, 2007. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2. *Science,* **316(5832)**, 1732–1735.

Sweeney, C., A. Karion, S. Wolter, T. Newberger, D. Guenther, J.A. Higgs, A. E. Andrews, P.M. Lang, D. Neff, E. Dlugokencky, J.B. Miller, S.A. Montzka, B.R. Miller, K.A. Masarie, S.C. Biraud, P.C. Novelli, M. Crotwell, A.M. Crotwell, K. Thoning, and P.P. Tans, 2015: Seasonal climatology of CO_2 across North America from aircraft measurements in the NOAA/ESRL Global Greenhouse Gas Reference Network. *J. Geophys. Res. Atmospheres*, **120**, 5155-5190. https://doi.org/10.1002/2014JD022591.

Tans, P.P., I.Y. Fung, and T. Takahashi, 1990: Observational constraints on the global atmospheric CO_2 budget. *Science*, **237**, 1431-1438.

Wofsy, S.C., Afshar, S., Allen, H., Apel, E., Asher, E.C., Barletta, B., et al., 2018: Atom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols. ORNL DAAC, Oak Ridge, TN, doi:10.3334/ORNLDAAC/1581.

Yang, M.M., J.D. Barrick, C. Sweeney, J.P. DiGangi, and J.R. Bennett. 2018. ACT-America: L1 Meteorological and Aircraft Navigational Data. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1574.