Geophysical Research Letters

RESEARCH LETTER

10.1029/2022GL100147

Special Section:
Carbon Weather: Toward the next generation of regional greenhouse gas inversion systems

Key Points:

- The seasonal amplitude of net ecosystem exchange (NEE) of CO_2 in the central and eastern temperate North America is underestimated in global atmospheric inversions
- The seasonal bias is not significantly different between inversions using OCO-2 v9 land nadir/glint observations and in situ observations
- The largest NEE biases are observed in U.S. croplands and eastern forests

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Y. Cui,
yqc5573@psu.edu

Citation:

Cui, Y. Y., Zhang, L., Jacobson, A. R., Johnson, M. S., Philip, S., Baker, D., et al. (2022). Evaluating global atmospheric inversions of terrestrial net ecosystem exchange CO_2 over North America on seasonal and sub-continental scales. *Geophysical Research Letters*, *49*, e2022GL100147. https://doi.org/10.1029/2022GL100147

Received 28 JUN 2022
Accepted 4 SEP 2022

Author Contributions:

Conceptualization: Yu Yan Cui, Kenneth J. Davis
Methodology: Yu Yan Cui, Li Zhang, Matthew S. Johnson, Kenneth J. Davis
Resources: Andrew R. Jacobson, Matthew S. Johnson, Sajeev Philip, David Baker, Frederic Chevallier, Andrew E. Schuh, Junjie Liu, Sean Crowell, Hélène E. Peiro, Feng Deng, Sourish Basu, Kenneth J. Davis
Software: Yu Yan Cui
Supervision: Kenneth J. Davis

Evaluating Global Atmospheric Inversions of Terrestrial Net Ecosystem Exchange CO_2 Over North America on Seasonal and Sub-Continental Scales

Yu Yan Cui[1], Li Zhang[1], Andrew R. Jacobson[2], Matthew S. Johnson[3], Sajeev Philip[4,5], David Baker[6], Frederic Chevallier[7], Andrew E. Schuh[6], Junjie Liu[8], Sean Crowell[9], Hélène E. Peiro[9], Feng Deng[10], Sourish Basu[11], and Kenneth J. Davis[1,12]

[1]Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA, USA, [2]University of Colorado Boulder and NOAA Earth System Research Laboratory, Boulder, CO, USA, [3]Earth Science Division, NASA Ames Research Center, Moffett Field, CA, USA, [4]NASA Ames Research Center, Universities Space Research Association, Mountain View, CA, USA, [5]Now at Centre for Atmospheric Sciences, Indian Institute of Technology Delhi, New Delhi, India, [6]Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA, [7]Laboratoire des Sciences du Climat et de L'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France, [8]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [9]School of Meteorology, University of Oklahoma, Norman, OK, USA, [10]Department of Physics, University of Toronto, Toronto, ON, Canada, [11]Goddard Space Flight Center, NASA GSFC GMAO and University of Maryland, Greenbelt, MD, USA, [12]Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA, USA

Abstract Atmospheric inversion estimates of net ecosystem exchange (NEE) of CO_2 are increasingly relevant to climate policy. We evaluated sub-continental, seasonal estimates of CO_2 NEE from nine global inversion systems that participated in the Orbiting Carbon Observatory-2 model intercomparison project (OCO-2 v9 MIP), using 98 research flights conducted over the central and eastern United States from 2016 to 2018 as part of the Atmospheric Carbon and Transport - America mission. We found that the seasonal amplitude of NEE in the central and eastern United States is underestimated in these models and model-data biases are largest for those inversions with the smallest seasonal flux amplitudes. These results were independent of whether the inversions used satellite or in situ data. The largest NEE biases were observed in the Midwest croplands and eastern forests. Future experiments are needed to determine the causes of the persistent biases and if they are associated with biases in annual flux estimates.

Plain Language Summary The exchange of CO_2 between terrestrial ecosystems and the atmosphere is an important component of the Earth's climate system. Atmospheric budgets are used to quantify this exchange globally, but these estimates are difficult to evaluate on a regional basis. We used a unique set of aircraft data to evaluate a set of state-of-the-science estimates of ecosystem-atmosphere CO_2 exchange in temperate North America. Nearly every estimate underestimated the seasonal amplitude of ecosystem-atmosphere CO_2 exchange (net photosynthesis too weak in the summer; respiration too weak in the winter) in this region. The source of atmospheric CO_2 data did not influence this finding. More study is needed to determine both the cause of these seasonal biases and the impact of this bias on annual net CO_2 flux estimates.

1. Introduction

Global atmospheric inversions are increasingly being used to estimate regional-scale net ecosystem exchange (NEE) of CO_2 (e.g., Byrne et al., 2020; Chen et al., 2021; Enting et al., 2012; Liu et al., 2017, 2021; Miller & Michalak, 2020; Palmer et al., 2019). National-scale comparisons of atmospheric inverse flux estimates to biomass inventories have shown promise, albeit with large uncertainty bounds in both products (e.g., Ciais et al., 2022). The global stocktake of the Paris Agreement (GST) has begun to consider global atmospheric inversions to inform national-scale CO_2 flux estimates (Chevallier, 2021; Ciais et al., 2022). Regional evaluation of the CO_2 NEE fluxes resulting from atmospheric inversions, however, is largely absent from the literature. Regional CO_2 NEE estimates from inversions must be evaluated rigorously at seasonal and regional resolution before they can be used with confidence to inform climate mitigation.

Writing – original draft: Yu Yan Cui, Kenneth J. Davis
Writing – review & editing: Yu Yan Cui, Matthew S. Johnson, David Baker, Frederic Chevallier, Kenneth J. Davis

The Orbiting Carbon Observatory-2 (OCO-2) Model Intercomparison Project (MIP) (Crowell et al., 2019) has brought together an ensemble of global atmospheric inversions systems. These inversion systems use a variety of atmospheric transport models, optimization algorithms, and prior flux estimates, but use common sets of atmospheric CO_2 data, including both column CO_2 (XCO_2) and in situ observations, to estimate NEE of CO_2 across the globe (Crowell et al., 2019; Peiro et al., 2022). These inversions arguably represent the most advanced set of global inverse estimates of NEE of CO_2, and they were evaluated with independent atmospheric observations but most of these evaluations have focused to date on global aggregate performance metrics for the annual budget (e.g., Chevallier et al., 2019; Crowell et al., 2019; Liu et al., 2021; Peiro et al., 2022). Currently, the evaluations of the inferred fluxes from most global inversion systems at the regional and seasonal scales are very limited.

The Atmospheric Carbon and Transport-America (ACT-America) project is a NASA Earth Venture Suborbital-2 mission designed to study the transport and fluxes of greenhouse gases in the midlatitudes (Davis et al., 2021). ACT-America campaigns took place in the central and eastern United States (US) during Summer 2016, Winter 2017, Fall 2017, Spring 2018, and Summer 2019 (Davis et al., 2021; Wei et al., 2021). These locations were chosen because they represent biologically productive middle-latitude ecoregions and experience vigorous synoptic weather that transports the CO_2 signature of those biological fluxes. Over 1,140 flight hours of data from 121 research flights distributed across the central and eastern United States sampled more than 30 synoptic sequences. Forty-five percent of the flight hours were within the atmospheric boundary layer (ABL). Temperate North America also has one of the densest in situ greenhouse gas monitoring networks in the world. The ACT data set thus provides a unique testbed to assess rigorously and for the first time the regional, seasonal performance of an array of global atmospheric CO_2 inversions.

We evaluated the seasonal, regional performance of the OCO-2 MIP flux inversions in temperate North America using the ACT observations. We compared observed ABL CO_2 mole fractions to corresponding CO_2 mole fractions simulated using OCO-2 MIP inversion products. We then quantified errors in the seasonal inverse estimates of CO_2 NEE. This work moves us toward quantitative understanding of the ability of global inversion systems to estimate regional NEE of CO_2.

2. Data and Methods

The model-data comparison framework used here follows the methods presented in Cui et al. (2021). This description focuses primarily on modifications to Cui et al. (2021)'s methods.

2.1. ACT-America Observations

ACT-America flights typically involved two aircraft in-flight patterns coordinated to sample atmospheric state throughout a portion of a synoptic weather system in either the MidAtlantic, MidWest or SouthCentral region of the United States (Davis et al., 2021). Both aircraft included highly calibrated observations of CO_2 (Baier et al., 2020; Wei et al., 2021). Each seasonal flight campaign lasted for 6 weeks, with 2 weeks focused on each of the three study regions. All stages of synoptic weather (prefrontal, frontal, and postfrontal) were sampled. Each flight included long-level flight legs within the ABL. The data sets, including the in situ aircraft observations (Davis et al., 2018) are documented and archived at Oak Ridge National Lab's Distributed Active Archive Center (https://daac.ornl.gov/actamerica).

2.2. CO_2 NEE Flux Inversion Products

The OCO-2 v9 MIP released a suite of gridded CO_2 flux inversions from 10 global inversion models encompassing the years 2015–2018. Details of the OCO-2 v9 MIP are described by Peiro et al. (2022) and Zhang et al. (2022) and in Table S1 of Supporting Information S1. The different inversion systems are standardized in the sense that they are required to assimilate the same four sets of atmospheric observations. The four observational data sources include the CO_2 mole fraction measurements from (a) in situ data ("IS") compiled in the GLOBALVIEW+ 5.0 (Cooperative Global Atmospheric Data Integration Project, 2019) and NRT v5.1 (CarbonTracker Team, 2019) ObsPack products; (b) the land nadir/land glint ("LNLG") retrievals of column-integrated CO_2 from OCO-2 v9; (c) OCO-2 ocean glint ("OG") v9 retrievals; and (d) a combination of the in situ and satellite data ("LNLGOGIS") (Kiel et al., 2019; Peiroet al., 2022). We evaluate 35 of these inversion products (nine inversion systems and four data experiments, and one inversion system did not conduct the LNLGOGIS experiment), and we sometimes limit our assessments to the LNLG and IS products from each of the nine inversion systems based on the suggestion from Cui et al. (2021) that the LNLG product is the best performing of the OCO-2 based inversion products in temperate North America and that all other inversions outperformed the OG-based inversion.

Prior flux estimates were only partly constrained in the MIP protocol. All models in the OCO-2 v9 MIP were required to use the same fossil fuel inventory from the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) 2019 version but were not limited to their choice of biospheric, oceanic, and fire prior fluxes. The selected prior flux inputs for the components of the biospheric, oceanic, and fire sources are listed in Table S2 of Supporting Information S1. Overall, there are seven different prior NEE of CO_2 estimates used in these inversion systems, six different prior estimates of the oceanic CO_2 fluxes, and four different prior fire CO_2 emissions estimates.

The OCO-2 v9 MIP produced CO_2 flux estimates with monthly resolution. Gridded global 3-hourly resolution CO_2 NEE products were generated for the four ACT-America campaign periods (i.e., summer 2016, winter 2017, fall 2017, and spring 2018) from the monthly flux products specifically for this study (Text S1 in Supporting Information S1).

2.3. Influence Functions

We estimated source-receptor relationships between CO_2 fluxes and atmospheric CO_2 mole fractions along the ACT-America flight tracks using Lagrangian particle dispersion modeling (e.g., Cui et al., 2021). In the study, we aggregated the ACT-America ABL CO_2 measurements in 10-min intervals, excluding take-off and landing portions. The ABL determination is described in Pal et al. (2020) and Davis et al. (2021). Each of the 10-min (roughly 60–70 km at typical flight speeds) intervals is treated as a receptor. We release 1,000 particles per receptor and simulate their backward transports for 10 days using FLEXPART v10.4 ("FLEXible PARTicle dispersion model," Pisso et al., 2019). The FLEXPART model was driven by ERA-interim reanalysis data (Dee et al., 2011).

2.4. Background Values

A portion of the simulated CO_2 mole fractions comes from the atmospheric state 10 days upwind of the observation point or receptor. We refer to this as the background value for atmospheric CO_2. The background values are determined using the CO_2 mole fraction fields simulated by each OCO-2 v9 MIP experiment using its posterior fluxes. We sampled the CO_2 mole fraction field at the locations in time and space when and where the particle trajectories' 10-day backward simulations terminated. The details of the background determination are described in Text S1 and Figure S1 of Supporting Information S1.

2.5. Model-Data Comparison Framework

We convolve each OCO-2 v9 MIP NEE flux estimate with the influence functions to simulate the biogenic contribution to the atmospheric CO_2 mole fractions along the ACT-America ABL flight tracks. We refer to this quantity as y_{modbio} (Cui et al., 2021). The contribution of ABL CO_2 mole fractions from fossil fuels, fire and ocean sources are calculated by convolving these surface flux maps with the 10-day influence functions. The NEE-related portion of the ACT-America CO_2 observations, which we refer to as y_{ACTbio} (Cui et al., 2021) are determined by subtracting the influence of fossil fuel, fire and ocean fluxes from the total ABL CO_2 measurements, as well as the determined background values (Section 2.4). We use the fossil fuel CO_2 emission estimates from the ODIAC, 2018 emission inventory (Oda et al., 2018), and fire emissions from the GFEDv4.1s wildfire emission inventory for all cases. The ocean CO_2 influence is derived from the monthly-averaged posterior oceanic CO_2 flux estimates from each experiment from the individual model of OCO-2 v9 MIP.

The flights were designed to sample ecosystem fluxes in the Central and Eastern US and maximize the influence of NEE on ABL CO_2. Numerical estimates in Cui et al. (2021) show that the fire and ocean fluxes have negligible contributions to the ABL mole fractions. Fossil fuel sources have a more significant, but moderate impact. We analyze the impact of the variability in boundary conditions across inversion systems in this study. We found that, along the flight's ABL track, differences in large-scale boundary conditions contributes less than 2 ppm to the spread of simulated CO_2 mole fractions, while the differences in fluxes within the domain lead to a spread of up

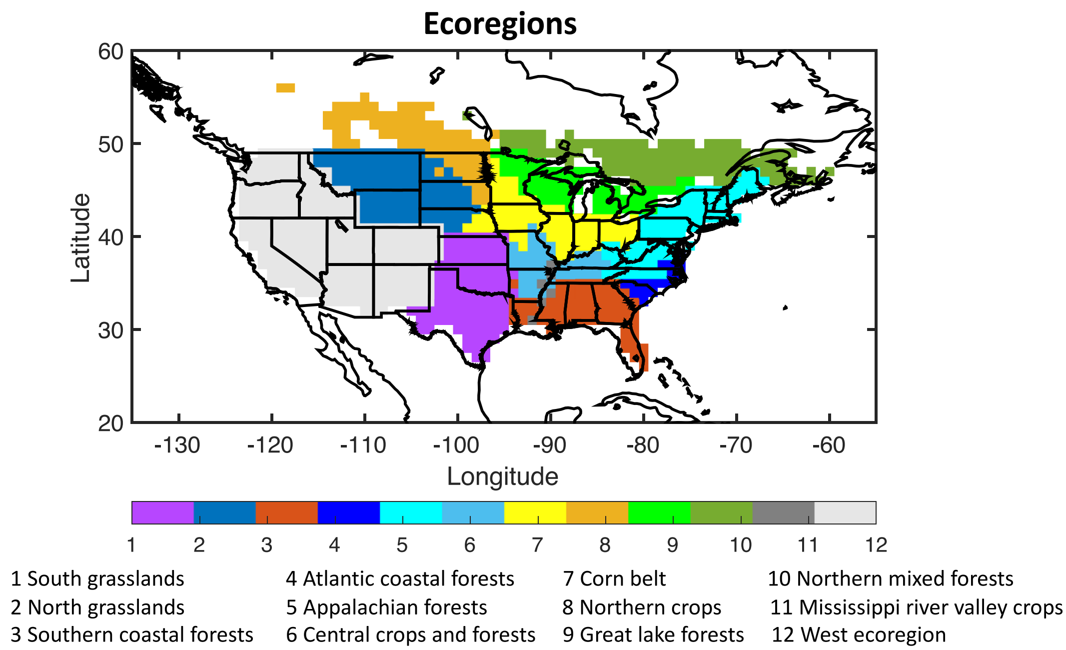

Figure 1. The spatial patterns of ecoregions in Temperate North America defined in the study. The influence functions of ACT-America flights cover most of the central and eastern temperate North America, and here we evaluate the mean bias error for the regions (1–9) that have the greatest influence on the ACT-America atmospheric boundary layer observations (Cui et al., 2021).

to 20 ppm (Figure S2 in Supporting Information S1). The large-scale transport uncertainty is small compared to the uncertainties caused by fluxes within the domain.

2.6. Evaluation Metrics

We evaluate the OCO-2 v9 MIP posterior NEE estimates using root-mean-square error (RMSE) and mean bias error (MBE) metrics. These metrics are computed by summing over all ABL observations (receptors) in each seasonal flight campaign, that is,

$$\mathrm{RMSE} = \frac{\sum_{i=1}^{N}\left(\sqrt{\left(y_{\mathrm{modbio}\_i} - y_{\mathrm{ACTbio}\_i}\right)^2}\right)}{N} \tag{1}$$

$$\mathrm{MBE} = \frac{\sum_{i=1}^{N}\left(y_{\mathrm{modbio}\_i} - y_{\mathrm{ACTbio}\_i}\right)}{N} \tag{2}$$

where i denotes each receptor, and N denotes the number of receptors within a seasonal flight campaign. In this way we evaluate how well the inversion systems simulate the spatial and temporal variability in regional fluxes (RMSE) and the mean seasonal flux magnitude (MBE).

2.7. Ecoregion-Based Evaluation Framework

To evaluate fluxes by ecoregion, we group the receptors by ecoregion and calculate the MBE values between the simulated and observed biological CO_2 mole fractions for each ecoregion and season. We attribute each receptor to the one eco-region which contributes the largest portion of the influence function for that receptor (Figure 1 and Figure S3 in Supporting Information S1).

3. Results

3.1. Regional Biases in NEE

The seasonal MBE for all members of OCO-2 v9 MIP are shown in Figure 2 and we find that nearly every member has the same sign of seasonal bias with respect to the ACT-America observations. The inverse estimates

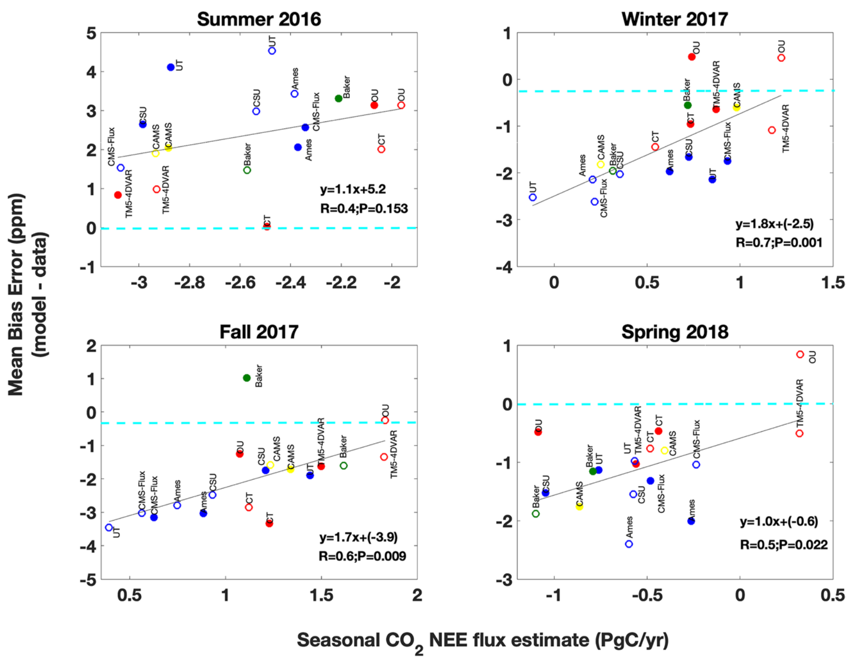

Figure 2. Seasonal net ecosystem exchange (NEE) of CO_2 in the Central and Eastern US (the domain is shown in Figure S4 of Supporting Information S1) as a function of the seasonal mean bias error (Equation 1). The panels represent the analysis for July and August 2016 ("Summer 2016"); February and March 2017 ("Winter 2017"); October and November 2017 ("Fall 2017"); and April and May 2018 ("Spring 2018"), respectively. Each OCO-2 v9 MIP member is labeled with the codes defined in (Table S1 in Supporting Information S1). The open circles denote the IS experiments, and the solid circles denote the land nadir/land glint experiments. The TM5 group (CT, OU, and TM5-4DVAR) is colored in red, the GEOS-Chem group (Ames, CMS-Flux, UT, and CSU) is colored in blue, the Baker model is in black, and the CAMS model is in yellow. The pink lines are linear regressions of all inversions for each season. The linear fits, correlation coefficients, and fractional probability that the quantities are not correlated are shown in the lower right of each panel.

of NEE of CO_2 yield positive MBE in the summer and negative MBE in the other seasons. These results suggest that nearly all inversions underestimate the magnitude of the seasonal cycle of NEE, with an underestimate of net photosynthesis in the summer and an underestimate of ecosystem respiration in the winter and fall. Spring is the one exception to this pattern, where the results suggest that the net photosynthesis is overestimated (NEE too negative). It is quite remarkable that this appears to be true for nearly every inversion system, independent of the source of data. The OU inversion system is one possible exception to this pattern.

The linear regression of all cases regardless of data source (Figure 2) establish the relationship between the bias errors determined with respect to the ACT-America flight data and the regionally-averaged, whole-season NEE estimates from the inversion models. We find that the MBEs are strongly correlated with the magnitude of seasonal NEE integrated across the entire study region, regardless of data source or inversion system. These correlations are statistically significant in all seasons save summer. The regionally, seasonally integrated inverse flux estimates with larger seasonal amplitudes have smaller MBE with respect to the ACT observations. The aircraft data are directly compared to a subsample of fluxes in space and time. It is possible that this subsample of fluxes in space and time might not be representative of the regionally, seasonally averaged flux estimates. This correlation shows that our aircraft-based evaluation does capture a pattern that is representative of the regional, seasonal fluxes.

Extrapolation of these correlations to zero MBE suggests that, on average, the MIP products underestimate dormant season respiration across the central and eastern US by about 1 PgC, with a similar magnitude but opposite sign underestimate (~-1 PgC) of net uptake in the summer season. Given that three of four seasons yield underestimates of positive NEE all about the same magnitude as the summer underestimates of net uptake, this suggests these inversion systems might yield annual NEE values for this region that are too negative, but that is a speculative inference given the limited temporal duration and discontinuous temporal coverage of the ACT-America flight observations.

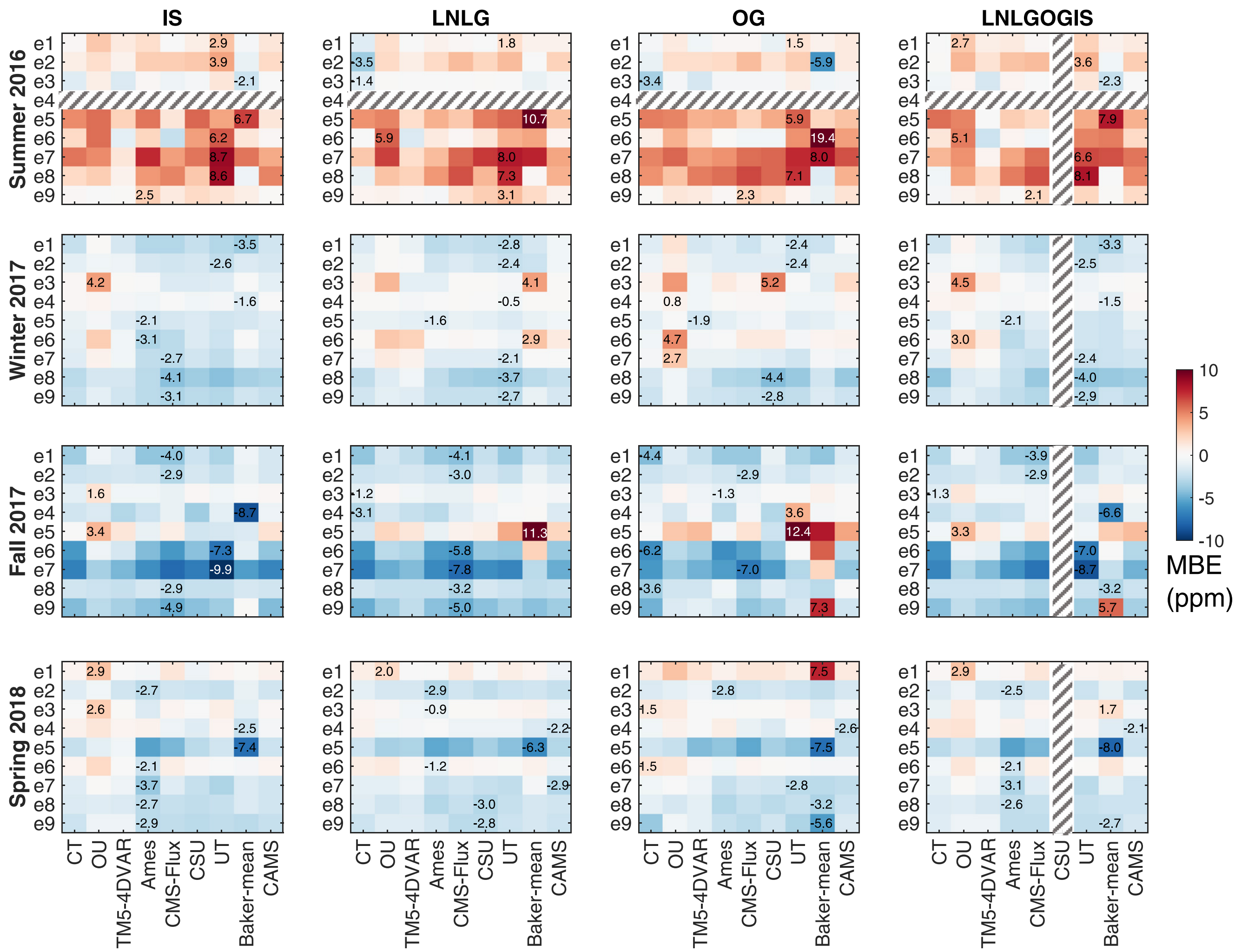

Figure 3. Mean Bias Error (MBE, ppm) for nine different ecoregions in Central and Eastern Temperate North America. The largest magnitude of MBE for each ecoregion and each season is written onto the cell. A warm color denotes a positive bias, and a cold color denotes a negative bias. The ecoregions are defined in Figure 1.

The inversion products from each model are only required to use the same fossil fuel emission and the same observational data sets, leaving many potential differences among the inversion systems including prior fluxes, atmospheric transport, and inversion algorithms. Given the limited similarity among the systems and the contrasting observations (IS and LNLG), the similarity in NEE performance relative to the ACT-America observations is striking. It is possible that a bias in our influence functions contributes to the MBE in Figure 2. The TM5 group shows the best performance among the transport models, with smaller MBEs (from −3.3 to 3.1 ppm) than the other transport models (from −3.5 to 4.5 ppm) across four seasons. Both our influence functions and the TM5 inversions are based on ERA reanalysis. This might explain the advantageous performance of the TM5 group of inversions in our analyses. Overall, the TM5-4DVAR model has the best performance across the different seasons.

3.2. Ecoregion-Based Bias Analyses

A number of broad patterns emerge when the MBE is evaluated for each ecoregion (Figure 3). In all seasons the patterns of ecoregion MBEs change relatively little as a function of the data source used in the inversion. Summer and fall have the largest overall MBEs. The largest MBEs are in the Appalachian forests (ecoregion #5), central crops and forest (ecoregion #6), the corn belt (ecoregion #7), and the northern grain belt (ecoregion #8). Appalachian forests (e5) stand out as having a positive fall MBE (overestimate of NEE in the inversions), unlike all other ecoregions during this season.

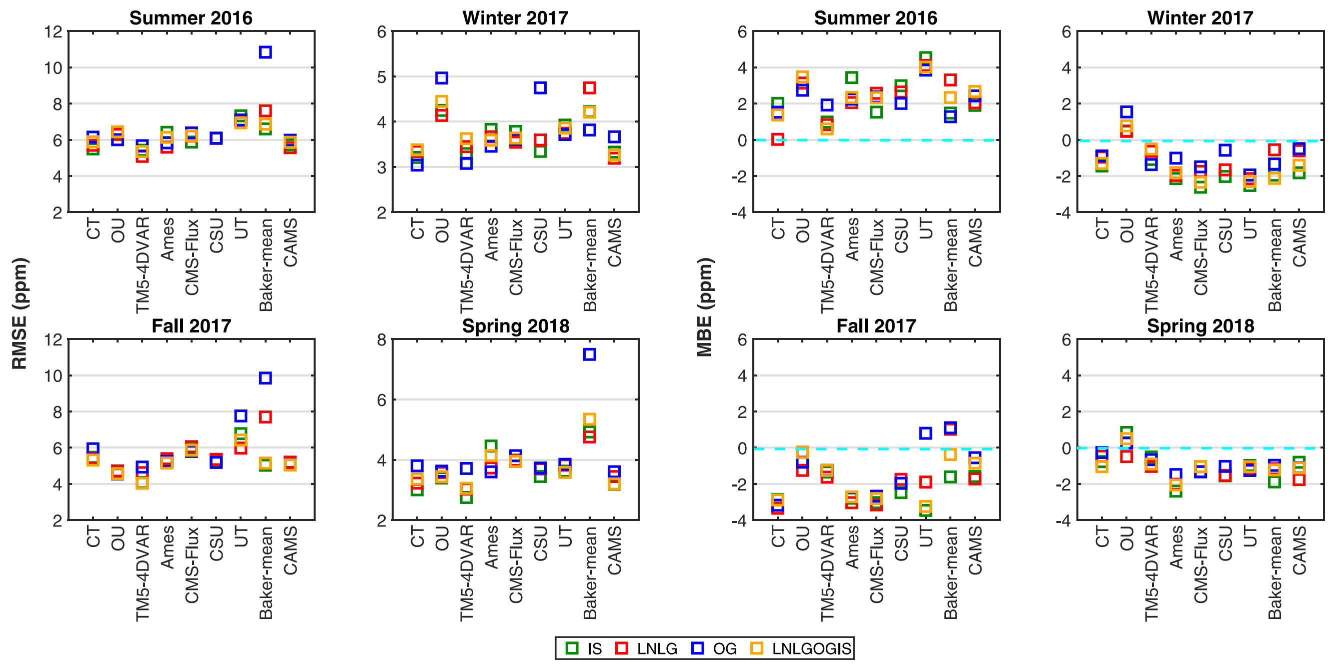

Figure 4. The root-mean-square error (RMSE) (left) and mean bias error (right) analysis of the posterior biogenic CO_2 computed from all inverse estimates of net ecosystem exchange of CO_2 compared to the observed atmospheric boundary layer CO_2 mole fractions from each of four seasonal ACT-America campaigns.

The performance of the inversion systems varies considerably when evaluated by ecoregion. Large MBEs of both signs are found in the Baker and UT models, which may imply less spatial correlation in the inversions used by these systems. The OU model MBEs most often diverge in sign from the other models during the dormant season, and the Ames and CMS-Flux models often have the largest negative MBEs in the dormant seasons, especially for the IS and LNLG inversions. The TM5-4DVAR model shows the smallest MBE across all ecoregions. The UT and Baker-mean models contain many of the peak positive biases across these ecoregions.

3.3. Regional RMSE and MBE Across Inversion Systems and Data Sources

A synthesis of RMSE and MBE results across all seasons, regions, inversions, and data sources (Figure 4) shows some patterns that are to be expected given the seasonal amplitude of NEE. Across all members of OCO-2 v9 MIP, spring and winter CO_2 NEE flux estimates have smaller RMSE levels than fall and summer estimates. These findings are roughly consistent with larger NEE, hence larger potential for model-data differences, in the more biologically active seasons.

The same synthesis (Figure 4) reveals some differentiation across inversion systems. Most of the models in the OCO-2 v9 MIP are not strongly sensitive to changes in the observational source. The Baker-mean model, in contrast, is relatively sensitive to the source data used in the inversion, especially to the OG experiment. The OU and CSU models are sensitive to the OG data during the wintertime as well. The UT model is sensitive to the different observing datasets during the fall months. This suggests that these inversion systems are the most data driven.

The MBE analysis as a function of the observational data set shows similar patterns to the RMSE analysis. MBE levels are smaller in winter and spring months than the fall and summer months, and the MBE level is smallest in the spring. During the fall months, the MBE levels for the CO_2 NEE flux estimates from the UT and Baker model still display large divergences across different observing datasets. The LNLGOGIS experiment includes both in situ and OCO-2 data but we do not find superior performance in the current global inversion system despite the increased data density.

4. Discussion

Our results show that the overall OCO-2 v9 MIP models underestimate the seasonal amplitude of NEE across central and eastern US ecosystems, regardless of data source. These results are consistent with the results of other studies using ACT-America observations (Cui et al., 2021; Feng et al., 2021; Zhang et al., 2022) but independent methods. Feng et al. (2021) finds similar results for the summer of 2016 (underestimate of net photosynthesis) but using an independent atmospheric transport system, the Weather Research and Forecast model (WRF), and only evaluating NEE from the CarbonTracker inversion system. Cui et al. (2021) used independent FLEXPART-WRF simulations associated with a Bayesian analysis to evaluate CarbonTracker inversions in the OCO2 v9 MIP at the sub-regions of Central and Eastern temperate North America. The seasonal mean bias analyses are in good agreement with both Cui et al. (2021) and this study for summer, fall, and spring, while an opposite sign of the mean bias estimates for winter likely due to the underestimates of ABL in WRF simulations of Cui et al. (2021). Zhang et al. (2022) co-sampled the ACT-America data in the posterior CO_2 fields of the OCO-2 v9 MIP. They found seasonal model-data mismatch patterns in boundary layer CO_2 from the OCO-2 v9 MIP which are largely consistent with an underestimate in seasonal NEE magnitudes. This suggests that our findings are not specific to the FLEXPART-ERA-interim atmospheric transport simulation used in this study. This does not rule out the possibility that all atmospheric transport systems available to us at this time might be biased.

Zhang et al. (2022) also found that the magnitudes of posterior, seasonally-and sectorally-averaged ABL-Free troposhere (FT) vertical CO_2 differences over the ACT-America flight domain were underestimated in the OCO-2 v9 MIP inversion systems, a result that could be consistent with either underestimated seasonal amplitudes of NEE, overestimated vertical mixing, or some combination of those factors. Zhang et al. (2022) evaluated mean wind speed and ABL depths in the inversion systems and did not find persistent biases. Cloud convective transport, however, could not be evaluated directly with ACT-America measurements.

Since the seasonal biases in NEE are of opposite sign, biases in annual NEE are not clear. The cause of these biases is also not clear given the analyses to date. The fact, however, that this seasonal bias appears to exist in nearly every inversion system, independent of data source, suggests an issue with the inversion systems used in the OCO-2 v9 MIP.

One possible explanation for this seasonal bias is that the prior NEE estimates used in the inversions have biased the posterior NEE estimates. Previous studies (e.g., Philip et al., 2019) have demonstrated the large impact of prior information on the current global inversion systems. Feng et al. (2021) showed that one model commonly used as a flux prior appears to underestimate the magnitude of summer NEE. We suggest further evaluation of the prior biospheric fluxes applied in the OCO-2 MIP.

The relationship between the bias errors determined by ACT-America data and the regional, seasonal NEE estimates from the models are not statistically significant in the summer. The different systems employ different inversion algorithms. Summer results might be more sensitive to inversion algorithms since fluxes are large and the corrections to the priors might be large in the summer. In addition, convective cloud transport is at a maximum in the summer and differences in atmospheric transport among inversion systems may increase in this season. Neither of these potential causes of divergence across inversion models would be clearly correlated with the regional, seasonal NEE estimates. The cause of this divergence in summer is not clear but is worthy of future study.

Atmospheric transport errors could also lead to persistent errors in either posterior fluxes or our inferences about these fluxes. We found that the TM5-based inversions had smaller seasonal NEE biases than the GEOS-Chem-based inversions. TM5 transport should be closely related to the ERA meteorology used to derive our influence functions. Schuh et al. (2019), suggested that TM5 mixes more vigorously in the vertical than does GEOS-Chem. This would lead to TM5-based inversions requiring stronger seasonal NEE of CO_2 to match ABL CO_2 observations since seasonal fluxes would be diluted within a larger atmospheric mixing volume. This appears consistent with our findings. These results do not say, however, which representation of atmospheric mixing is more realistic.

It is difficult from this suite of results to conclude which fluxes are most accurate because of the compensating influences of flux and mixing on atmospheric CO_2, but these analyses do put boundaries on the problem. Likely biases appear to tend toward either an overestimate of mixing in TM5 and/or an underestimate of the seasonal amplitude of NEE, particularly in GEOS-Chem-based inversions, with biases in both systems leading to the underestimate of the ABL-FT CO_2 differences documented by Zhang et al. (2022). Overestimates in the seasonal amplitude of NEE in the inversion systems appear unlikely to be consistent with this suite of results.

Schuh et al. (2019) showed that globally, the differences in atmospheric mixing between these two systems led to large differences in inverse estimates of annual NEE of CO_2, emphasizing the need to identify and minimize the biases we have documented.

5. Conclusions

The persistence of the seasonal underestimate of the magnitude of seasonal NEE in the central and eastern United States across all inversion systems and data sources suggests a potentially important bias within our most advanced atmospheric inversion systems. This pattern of results deserves further investigation given the growing importance of atmospheric inversions in global climate policy and the potential impact on annual NEE estimates. The lack of sensitivity of this result to the source of atmospheric CO_2 observations shows either encouraging consistency across observations or limited impact of the observations on the inversion products. The difference in seasonal biases in NEE of CO_2 as a function of the atmospheric transport model underscores the continued importance of evaluating the accuracy of the transport models used in these inversions. Experimentation with the atmospheric inversion systems is needed to explore the causes of these biases. The ACT-America flight observations, merging regionally-, seasonally-representative mole fraction data and direct observations of important atmospheric transport metrics, serve as a test bed for continued development of atmospheric inversion systems.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All ACT-America in situ data used in the manuscript can be found at the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL-DAAC) (https://doi.org/10.3334/ORNLDAAC/1556). The FLEXPART v10.4 model can be found online (https://www.flexpart.eu/wiki/FpInstall).

Acknowledgments
The ACT-America project is a NASA Earth Venture Suborbital 2 project funded by NASA's Earth Science Division. The authors would like to acknowledge the NASA grants: NNX15AG76G to Penn State and NNX15AJ07G to Colorado State, and the support for the OCO-2 v9 flux model inter-comparison project provided through NASA grant #80 NSSC 18K0909, and NASA's Carbon Monitoring System (80NSSC21K1060). MJ acknowledges the internal funding from NASA's Earth Science Research and Analysis Program. The authors also would like to acknowledge NASA's Earth System Science Pathfinder Program Office, NASA's Airborne Sciences Program, NASA's Atmospheric Science Data Center, and NASA's Pleiades supercomputing facilities.

References

Baier, B. C., Sweeney, C., Choi, Y., Davis, K. J., DiGangi, J. P., Feng, S., et al. (2020). Multispecies assessment of factors influencing regional CO_2 and CH4 enhancements during the winter 2017 ACT-America campaign. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD031339. https://doi.org/10.1029/2019JD031339

Byrne, B., Liu, J., Lee, M., Baker, I., Bowman, K. W., Deutscher, N. M., et al. (2020). Improved constraints on Northern Extratropical CO_2 fluxes obtained by combining surface-based and space-based atmospheric CO_2 measurements. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD032029. https://doi.org/10.1029/2019JD032029

CarbonTracker Team. (2019). Retrieved from https://gml.noaa.gov/ccgg/carbontracker/CT-NRT.v2019-2/

Chen, Z., Liu, J., Henze, D. K., Huntzinger, D. N., Wells, K. C., Sitch, S., et al. (2021). Linking global terrestrial CO_2 fluxes and environmental drivers: Inferences from the Orbiting Carbon Observatory 2 satellite and terrestrial biospheric models. *Atmospheric Chemistry and Physics*, *21*(9), 6663–6680. https://doi.org/10.5194/acp-21-6663-2021

Chevallier, F. (2021). Fluxes of carbon dioxide from managed ecosystems estimated by national inventories compared to atmospheric inverse modelling. *Geophysical Research Letters*, *48*, e2021GL093565. https://doi.org/10.1029/2021GL093565

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., & Cozic, A. (2019). Objective evaluation of surface-and satellite-driven carbon dioxide atmospheric inversions. *Atmospheric Chemistry and Physics*, *19*(22), 14233–14251. https://doi.org/10.5194/acp-19-14233-2019

Ciais, P., Bastos, A., Chevallier, F., Lauerwald, R., Poulter, B., Canadell, J. G., et al. (2022). Definitions and methods to estimate regional land carbon fluxes for the second phase of the REgional Carbon Cycle Assessment and Processes Project (RECCAP-2). *Geoscientific Model Development*, *15*(3), 1289–1316. https://doi.org/10.5194/gmd-15-1289-2022

Cooperative Global Atmospheric Data Integration Project. (2019). *Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2018*. NOAA Earth System Research Laboratory, Global Monitoring Division. https://doi.org/10.25925/20190812

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., et al. (2019). The 2015-2016 carbon cycle as seen from OCO-2 and the global in situ network. *Atmospheric Chemistry and Physics*, *19*(15), 9797–9831. https://doi.org/10.5194/acp-19-9797-2019

Cui, Y. Y., Jacobson, A. R., Feng, S., Wesloh, D., Barkley, Z. R., Zhang, L., et al. (2021). Evaluation of CarbonTracker's inverse estimates of North American net ecosystem exchange of CO_2 from different observing systems using ACT-America airborne observations. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD034406. https://doi.org/10.1029/2020JD034406

Davis, K., Obland, M., Lin, B., Lauvaux, T., O'Dell, C., Meadows, B., & Pauly, R. (2018). *Act-America: L3 merged in situ atmospheric trace gases and flask data*. ORNL DAACm. https://doi.org/10.3334/ORNLDAAC/1593

Davis, K. J., Browell, E. V., Feng, S., Lauvaux, T., Obland, M. D., Pal, S., et al. (2021). The atmospheric carbon and transport (ACT)—America mission. *Bulletin of the American Meteorological Society*, *102*(9), 1–54. https://doi.org/10.1175/bams-d-20-0300.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Enting, I. G., Rayner, P. J., & Ciais, P. (2012). Carbon cycle uncertainty in REgional Carbon Cycle Assessment and Processes (RECCAP). *Biogeosciences*, *9*(8), 2889–2904. https://doi.org/10.5194/bg-9-2889-2012

Feng, S., Lauvaux, T., Williams, C. A., Davis, K. J., Zhou, Y., Baker, I., et al. (2021). Joint CO_2 mole fraction and flux analysis confirms missing processes in CASA terrestrial carbon uptake over North America. *Global Biogeochemical Cycles*, *35*(7), e2020GB006914. https://doi.org/10.1029/2020GB006914

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. *Atmospheric Measurement Techniques*, *12*(4), 2241–2259. https://doi.org/10.5194/amt-12-2241-2019

Liu, J., Baskaran, L., Bowman, K., Schimel, D., Bloom, A. A., Parazoo, N. C., et al. (2021). Carbon monitoring system flux net biosphere exchange 2020 (CMS-Flux NBE 2020). *Earth System Science Data*, *13*(2), 299–330. https://doi.org/10.5194/essd-13-299-2021

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

Miller, S. M., & Michalak, A. M. (2020). The impact of improved satellite retrievals on estimates of biospheric carbon balance. *Atmospheric Chemistry and Physics*, *20*(1), 323–331. https://doi.org/10.5194/acp-20-323-2020

Oda, T., Maksyutov, S., & Andres, R. J. (2018). The open-source data inventory for anthropogenic CO_2, version 2016 (2016): A global monthly fossil fuel CO_2 gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data*, *10*(1), 87–107. https://doi.org/10.5194/essd-10-87-2018

ODIAC. (2018). Retrieved from https://gmao.gsfc.nasa.gov/gmaoftp/sourish/ODIAC/2018/distrib/

Pal, S., Davis, K. J., Lauvaux, T., Browell, E. V., Gaudet, B. J., Stauffer, D. R., et al. (2020). Observations of Greenhouse gas changes across summer frontal boundaries in the Eastern United States. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD030526. https://doi.org/10.1029/2019JD030526

Palmer, P. I., Feng, L., Baker, D., Chevallier, F., Bösch, H., & Somkuti, P. (2019). Net carbon emissions from African biosphere dominate pan-tropical atmospheric CO_2 signal. *Nature Communications*, *10*(1), 3344. https://doi.org/10.1038/s41467-019-11097-w

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., et al. (2022). Four years of global carbon cycle observed from the Orbiting Carbon Observatory 2 (OCO-2) version 9 and in situ data and comparison to OCO-2 version 7. *Atmospheric Chemistry and Physics*, *22*, 1097–1130. https://doi.org/10.5194/acp-22-1097-2022

Philip, S., Johnson, M. S., Potter, C., Genovesse, V., Baker, D. F., Haynes, K. D., et al. (2019). Prior biosphere model impact on global terrestrial CO_2 fluxes estimated from OCO-2 retrievals. *Atmospheric Chemistry and Physics*, *19*(20), 13267–13287. https://doi.org/10.5194/acp-19-13267-2019

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., Cassiani, M., Eckhardt, S., et al. (2019). The Lagrangian particle dispersion model FLEXPART version 10.4. *Geoscientific Model Development*, *12*, 4955–4997. https://doi.org/10.5194/gmd-12-4955-2019

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., et al. (2019). Quantifying the impact of atmospheric transport uncertainty on CO_2 surface flux estimates. *Global Biogeochemical Cycles*, *33*(4), 484–500. https://doi.org/10.1029/2018GB006086

Wei, Y., Shrestha, R., Pal, S., Gerken, T., Feng, S., McNelis, J., et al. (2021). Atmospheric Carbon and Transport—America (ACT-America) data sets: Description, management, and delivery. *Earth and Space Science*, *8*(7), e2020EA001634. https://doi.org/10.1029/2020EA001634

Zhang, L., Davis, K. J., Schuh, A. E., Jacobson, A. R., Pal, S., Cui, Y. Y., et al. (2022). Multi-season evaluation of CO_2 weather in OCO-2 MIP models. *Journal of Geophysical Research: Atmospheres*, *127*, e2021JD035457. https://doi.org/10.1029/2021JD035457