Seasonal strength of terrestrial net ecosystem CO2 exchange from North America is underestimated in global inverse modeling

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

[1]The Pennsylvania State University
[2]NOAA/GML/CIRES
[3]NASA Ames Research Center
[4]Colorado State University
[5]Laboratoire des Sciences du Climat et de l'Environnement (LSCE)
[6]Jet Propulsion Laboratory
[7]University of Oklahoma
[8]University of Toronto
[9]NASA GSFC GMAO / University of Maryland

November 22, 2022

Abstract

We evaluate terrestrial net ecosystem-atmosphere exchange (NEE) of CO2 from nine global inversion systems that inferred fluxes from four CO2 observational sources. We use 98 flights in the central and eastern U.S. from the ACT-America aircraft mission to conduct this sub-continental, seasonal-scale evaluation. We use Lagrangian particle dispersion modeling (FLEXPARTv10.4-ERA-Interim) to compare observed and simulated regional biogenic CO2 mole fractions. We find a positive bias (modeled CO2 > observed) in the summer and negative bias (modeled CO2 < observed) in dormant seasons across most flux products, suggesting that the seasonal strength of CO2 NEE is underestimated in these inverse models. Fluxes inferred from OCO-2 v9 satellite land nadir/glint observations yield an error level that is similar to fluxes inferred from in-situ data. Large bias errors are observed in the croplands and eastern forests. Future experiments are needed to determine if these seasonal biases are associated with biases in net annual flux estimates.

Seasonal strength of terrestrial net ecosystem CO_2 exchange from North America is underestimated in global inverse modeling

Yu Yan Cui[1], Li Zhang[1], Andrew R. Jacobson[2,3], Matthew S. Johnson[4], Sajeev Philip[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]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[3]NOAA Earth System Research Laboratory, Global Monitoring Laboratory, Boulder, CO, USA
[4]Earth Science Division, NASA Ames Research Center, Moffett Field, CA, USA
[5]Universities Space Research Association, NASA Ames Research Center, Mountain View, CA, USA
[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]University of Oklahoma, School of Meteorology, Norman, OK, USA
[10]Department of Physics, University of Toronto, Toronto, ON, Canada
[11]NASA GSFC GMAO / University of Maryland, Goddard Space Flight Center, Greenbelt, MD, USA
[12]Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA, USA

Key Points:

- The seasonal amplitude of CO_2 NEE in the central and eastern US is underestimated in most global inversion models.
- This season bias is not significantly different between inversions using OCO-2 v9 LNLG and in situ observations.
- Largest CO_2 flux biases are observed in U.S. croplands and eastern forests.

Corresponding author: Yu Yan Cui, yqc5573@psu.edu

Abstract

We evaluate terrestrial net ecosystem-atmosphere exchange (NEE) of CO_2 from nine global inversion systems that inferred fluxes from four CO_2 observational sources. We use 98 flights in the central and eastern U.S. from the ACT-America aircraft mission to conduct this sub-continental, seasonal-scale evaluation. We use Lagrangian particle dispersion modeling (FLEXPARTv10.4-ERA-Interim) to compare observed and simulated regional biogenic CO_2 mole fractions. We find a positive bias (modeled CO_2 > observed) in the summer and negative bias (modeled CO_2 < observed) in dormant seasons across most flux products, suggesting that the seasonal strength of CO_2 NEE is underestimated in these inverse models. Fluxes inferred from OCO-2 v9 satellite land nadir/glint observations yield an error level that is similar to fluxes inferred from in-situ data. Large bias errors are observed in the croplands and eastern forests. Future experiments are needed to determine if these seasonal biases are associated with biases in net annual flux estimates.

Plain Language Summary

The quantification of terrestrial net ecosystem-atmosphere exchange (NEE) of CO_2 is important to our understanding of the carbon cycle and constitutes an important contribution to the science which underpins climate policy. We use multi-season aircraft observations to evaluate the estimates of seasonal, regional NEE of CO_2 derived from both satellite and ground-based observations of atmospheric CO_2 using nine different global data analysis systems. Our analysis focuses on terrestrial ecosystems in the central and eastern United States. We find that nearly every analysis model yields an underestimate of the seasonal strength of NEE of CO_2 (net photosynthesis too weak in the summer; respiration too weak in the winter) regardless of the CO_2 data source. Additional 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

Accurate, spatially- and temporally- resolved carbon flux estimation is essential for improving climate projections and informing carbon management and policy (e.g., Arora et al., 2020; Millar et al., 2017). A thorough knowledge of the biological CO_2 fluxes from a variety of ecosystems across different geographic locations facilitates total carbon flux estimation and the establishment of national and state implementation plans (e.g., Pan et al., 2011; Tan et al., 2015; J. B. Miller et al., 2020; Wang et al., 2020)(California's Natural and Working Lands (NWL) Implementation Plan:`https://ww2.arb.ca.gov/our-work/programs/natural-and-working-lands`) . Ecosystem carbon-stock inventories and terrestrial biogeochemical models are commonly used to provide biospheric carbon fluxes for policy planning (e.g., Tan et al., 2015)(California's NWL Inventory:`https://ww2.arb.ca.gov/nwl-inventory`). Atmospheric inversion of CO_2 mole fraction observations to estimate biospheric CO_2 fluxes is an important and complementary avenue for independent evaluation of ecosystem carbon flux estimates (e.g., Ciais et al., 2010; Chevallier, 2021). These methods have benefited from the expansion of long-term atmospheric observing systems including both ground-based, airborne and space-based platforms (Crisp et al., 2008; Andrews et al., 2014; Sweeney et al., 2015; Karion et al., 2020).

An atmospheric inversion of CO_2 mole fractions optimizes CO_2 fluxes in such a way that simulated atmospheric CO_2 mole fractions agree better with observations (e.g., Rayner et al., 2019). Gridded global CO_2 fluxes are available from several multi-year atmospheric inversions, many of which are frequently updated to quantify CO_2 surface fluxes (e.g., CarbonTracker, `https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/` or the Copernicus Atmosphere Monitoring Service, `https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion`). The inversion models use prior

CO_2 flux estimates of different source components, including fossil fuel, biosphere, fire, and ocean. In general, most global inversion systems optimize the magnitude of land biospheric and oceanic CO_2 flux terms while leaving fossil fuel emissions "fixed" to derive the optimal solution.

These CO_2 flux inversions estimate fluxes across the globe with a variety of spatial resolutions. Accurate regional flux information has the potential to inform policy planning and carbon management. To date, regional flux estimates within global inversions have shown large differences (Peylin et al., 2013; Crowell et al., 2019). Rigorous evaluation of current CO_2 flux inversion products in time and space is needed to improve atmospheric inversions to the point of being a sound, verified source of information to be used in regional carbon accounting.

Aircraft field campaigns are well-suited for regional flux evaluation. Aircraft field campaigns have been deployed in many different regions to investigate CO_2 NEE surface fluxes, including the CO_2 Budget and Rectification Airborne study over temperate North America (COBRA) (Gerbig et al., 2003), the Arctic-Boreal Vulnerability Experiment in boreal North America (ABoVE) (C. E. Miller et al., 2019), and the Atmospheric Carbon and Transport-America Earth Venture Suborbital mission (ACT-America)(Davis et al., 2021). Several studies have been conducted to evaluate the global CO_2 flux inversions using independent aircraft CO_2 measurements above the atmospheric boundary layer (ABL) and focus on a large domain, such as global or continental scale (Liu & Bowman, 2016; Chevallier et al., 2019; Gaubert et al., 2019; Liu et al., 2021). To date, few studies have been conducted to evaluate the seasonal and sub-continental estimates of the global CO_2 flux inversions. ACT-America is the largest carbon-centric aircraft mission conducted in any midlatitude, continental environment. The multi-seasonal ACT-America campaigns were held in the central and eastern United States (U.S.) during Summer 2016, Winter 2017, Fall 2017, Spring 2018, and Summer 2019 (Davis et al., 2021; Wei et al., 2021). Over 1140 flight hours of data, roughly 45% of which were within ABL, were collected over the course of 121 research flights distributed across the central and eastern United States. The ACT-America flights sampled CO_2 mole fractions from the ABL to the upper free troposphere and were oriented to capture synoptic weather passages typical of each season and region (Pal et al., 2020; Wei et al., 2021). This multi-seasonal weather-oriented aircraft campaign provides a unique opportunity to assess inverse estimates of regional CO_2 NEE.

Global CO_2 flux inversions can be based on ground-based CO_2 monitoring or satellite-based retrievals of the total column CO_2 (XCO_2) mole fractions. These observing systems provide complementary temporal and spatial representativeness. The Orbiting Carbon Observatory-2 (OCO-2) satellite was launched in July 2014 and was designed to quantify sources and sinks of CO_2 across the globe (Eldering et al., 2017). The OCO-2 v9 model intercomparison project (MIP)(Peiro et al., 2021) produced a suite of multiyear (2015-2019) gridded global CO_2 flux inversion products, including the NEE of CO_2. The OCO-2 v9 MIP includes 10 global CO_2 data assimilation systems and is designed to assimilate both CO_2 in-situ data and the OCO-2 v9 column CO_2 data individually or collectively. We take advantage of the large spatial coverage and multi-seasonal sampling of ACT-America to evaluate the OCO-2 v9 MIP CO_2 NEE of temperate North America by comparing observed ABL CO_2 mole fractions to the corresponding simulated CO_2 mole fractions using the series of OCO-2 v9 MIP CO_2 flux inversion products. We apply two evaluation metrics to quantify the errors in CO_2 NEE from commonly-used global CO_2 inversion systems (applied in the OCO-2 v9 MIP) with respect to the independent airborne observations at sub-continental and ecoregional scales. The results are presented in Section 3, after the description of our data and methods in Section 2. The discussions and conclusion are shown in Section 4.

2 Data and methods

2.1 CO_2 NEE flux inversion products

OCO-2 v9 MIP released a suite of ten gridded CO_2 flux inversion products at the global scale encompassing the years 2015-2018. The different inversion systems are standardized in the sense that they are required to assimilate the same four sets of atmospheric observations and use the same fossil fuel CO_2 emissions as part of the inversion system inputs. The ten global CO_2 data assimilation systems are described by Peiro et al. (2021); Zhang et al. (2021) and some additional information are given in Text S1. The four observational data sources include the CO_2 mole fraction measurements from 1) 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; 2) the land nadir/land glint (LNLG) retrievals of column-integrated CO_2 from OCO-2 v9; 3) OCO-2 ocean glint (OG) v9 retrievals; and 4) a combination of the in situ and satellite data (LNLGOGIS). The suite of multiyear gridded CO_2 flux inversions are the monthly averaged products (`https://gml.noaa.gov/ccgg/OCO2_v9mip/`). In this study, ancillary gridded global CO_2 NEE products at 3-hourly resolution from nine members of OCO-2 v9 MIP (Text S1) was created for the four ACT-America Campaign periods (summer 2016, winter 2017, fall 2017, and spring 2018). All models in OCO_2 v9 MIP were required to use the same fossil fuel inventory from the Open-source Data Inventory for Anthropogenic CO_2 (ODIAC) 2018 version but were not limited in their choice of biospheric, oceanic and fire prior fluxes. The prior flux inputs for the components of the biospheric, oceanic, and fire sources are listed in Table S2. Overall, there are 7 different prior NEE of CO_2 estimates used in these inversion systems, 6 different prior estimates of the oceanic CO_2 fluxes, and 4 different prior fire CO_2 emissions estimates.

2.2 Influence functions

We established the source-receptor relationship between CO_2 NEE fluxes and atmospheric CO_2 enhancement/depletion along flight tracks using the Lagrangian particle dispersion modeling technique (e.g., Cui et al., 2021). In the study, we aggregated the ACT-America CO_2 measurements in the ABL, excluding take-off and landing portions, to the 10-minute intervals to match the spatial resolution of the transport simulations in the global inversion systems. The ABL determination is described in Pal et al. (2020) and Davis et al. (2021). Each of the 10-minute (roughly 60-70 km at typical flight speeds) intervals is treated as a receptor and we release 1000 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 the ERA-interim reanalysis data (0.75 x 0.75 degree, 6-hourly).

2.3 Background values

To determine the background values, 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 CO_2 mole fraction fields are from the long-term forward simulation from each OCO-2 v9 MIP model within the optimized fluxes from each experiment. The total number of the CO_2 mole fraction fields used here are 35 (9 models and 4 experiments, and the CSU model did not implement the LNLGOGIS experiment).

Specifically, we use the option of FLEXPART to output the spatially and temporally resolved sensitivity field (dimensionless and the range is from 0 to 1) of each receptor used in the study to the initial conditions, interface with the CO_2 mole fraction fields when and where particles are terminated to determine the background value for each receptor (Text S1 and Figure S1).

2.4 Evaluation metrics

We convolve each CO_2 NEE flux product to the atmospheric mole fractions along the ACT-America ABL flight tracks and compare them with the enhancement or depletion levels of the NEE-related CO_2 mole fractions within the ABL observed by ACT-America. The enhancement/depletion levels of the CO_2 mole fractions sampled by ACT-America flights are total CO_2 influenced by different CO_2 sources. The influence of biological sources dominates the aircraft data because the flights were designed to fly over the ecosystems in the Central and Eastern US. We obtain the enhancement/depletion levels of the NEE- related CO_2 mole fractions along flights after extracting the portions influenced by the fossil fuels, fire and ocean from the total CO_2 measurements, as well as the determined regional background values described in Section 2.3 (Cui et al., 2021). The influences of fossil fuels, fire and ocean are calculated using the influence function to convolve their surface fluxes within the 10-day span. We use the fossil fuel CO_2 emission estimates from the ODIAC 2018 emission inventory, and fire emissions from the GFED v4.1s wildfire emission inventory for all cases. The ocean CO_2 influence is derived from the convolution of the influence function and the monthly-averaged posterior oceanic CO_2 flux estimates from each experiment of the individual model in OCO-2 v9 MIP. In the study, we only used the boundary-layer CO_2 mole fractions of the ACT-America flights in the evaluation. Numerical estimates in Cui et al. (2021) show that the fire and ocean fluxes have very small contributions to the ABL mole fractions. Fossil fuel sources have a more significant, but moderate impact.

Cui et al. (2021) used the root-mean-square error (RMSE) metric (equation 2) to evaluate inversion products of the CarbonTracker model, one of OCO-2 v9 MIP ensemble members, based on the comparisons between the simulated and ACT-America referenced NEE-related CO_2 mole fractions. In this study, we apply the RMSE metric to nine models of OCO-2 v9 MIP. Furthermore, we focus more on the mean bias error (MBE) metric analysis (equation 3) in the CO_2 mole fraction space to investigate the bias error of each inversion case in OCO-2 v9 MIP.

$$RMSE = \frac{\sum_{i=1}^{N} \sqrt{(y_{modbio_i} - y_{ACTbio_i})^2}}{N} \tag{1}$$

$$MBE = \frac{\sum_{i=1}^{N} (y_{modbio_i} - y_{ACTbio_i})}{N} \tag{2}$$

, where i denotes each receptor, and N denotes the number of receptors. More details of y_{modbio_i} and y_{ACTbio_i} are described in Cui et al. (2021). Similar to Cui et al. (2021), our evaluation is seasonal. The RMSE and MBE values are calculated for each campaign (i.e each season). The flux product associated with smaller RMSE values indicates better spatially and temporally resolved flux estimates. The MBE analysis is also applied for each campaign. Smaller biases imply NEE of CO_2 that is most consistent with the mean impact of biogenic fluxes on ABL CO_2.

2.5 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 group. The ecoregion-based MBE analysis are subsets of the overall MBE analysis. We present the "zoom-in" maps to investigate the spatial origins of the MBE values and show the maximum MBE value for each ecoregion associated with the corresponding inversion case.

We attribute the receptors along the flight tracks to different ecoregions, taking advantage of the source-receptor relationship obtained from the Lagragian framework. Specif-

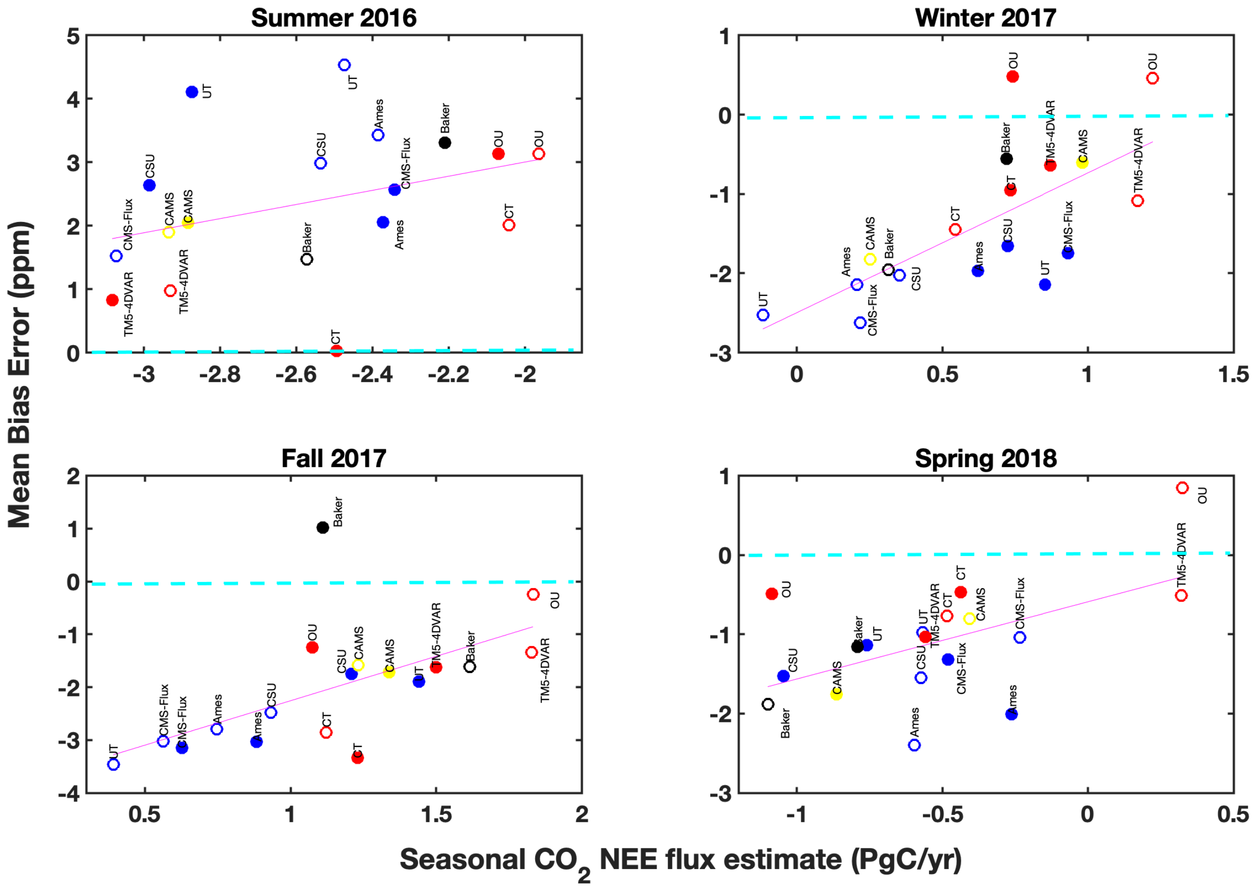

Figure 1. Seasonal NEE of CO_2 estimated from OCO-2 v9 MIP in Central and Eastern US (Text S1 and Figure S1) as a function of seasonal Mean Bias Error (MBE) values in posterior fluxes from the OCO- 2 v9 -MIP calculated using ACT-America ABL CO_2 mole fraction observations and a Lagrangian particle dispersion model (Cui et al., 2021). The observations and calculated NEE of CO_2 encompass July-August 2016 ("Summer 2016"); February-March 2017 ("Winter 2017"); October-November 2017 ("Fall 2017"); and April-May 2018 ("Spring 2018"). The open circles denote the IS experiments, and the solid circles denote the LNLG 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 cases for each season.

ically, we attribute each receptor to one eco-region which contributes the largest influence function for that receptor (Text S1 and Figure S2). We group the segments of CO_2 mole fractions along the flight tracks into different ecoregions and apply the MBE analysis for each group to investigate the associated seasonal bias levels aligned with the ecoregion regions of the temperate North America area. The overall spatial coverages of the influence functions of ACT-America are shown in Cui et al. (2021). We focus on region 1-9 in this study, which contribute largest influence on the enhancement/depletion of CO_2 mole fractions along ACT-America ABL flight tracks.

3 Results

Figure 1 shows seasonal Mean Bias Error (MBE) levels to the seasonal NEE estimation of OCO-2 v9 MIP members. We focus here on the flux estimates from the in-situ ("IS") and the OCO-2 v9 land nadir/land glint ("LNLG") experiments, which Cui et al., (2021) suggests are the most reliable NEE estimates for the central and eastern US. We find correlations between OCO-2 v9 MIP seasonal NEE estimates and seasonal MBE. The corresponding correlation coefficient (p-value) to the four campaigns are 0.4 ($p=0.15$), 0.7 ($p=0.001$), 0.6 ($p=0.009$), and 0.5 ($p=0.02$), respectively. The correlations are statistically significant for the winter, fall and spring months. Figure 1 shows that posterior estimates of NEE of CO_2 are underestimated in the IS and LNLG experiments compared to observations during winter, fall, and spring. Posterior estimates of NEE of CO_2 are overestimated (not sufficiently negative) during the summer. The TM5-4DVAR and OU models have the best performance during winter and fall seasons. The TM5-4DVAR and CT model within the LNLG experiment have the best performance during the summer.

The inversion products from each model are only required to use the same fossil fuel emission and the same observational datasets, leaving many potential differences among the inversion systems including prior fluxes, transport, and inversion algorithms. Therefore, some of the performance differences of the inversion systems is caused by the differences of these model framework components, enabling limited diagnosis of the causes of the MBEs. Overall, the TM5-4DVAR model has the best performance across the different seasons. The TM5 group shows the best performance among the transport models, with smaller MBEs than the other transport models across four seasons. The OCO-2 v9 land nadir/land glint experiment yields the MBE level that is similar to, or better (e.g winter) than, the in situ data experiment. We have used one transport model to create the influence functions used to link NEE of CO_2 to ABL CO_2 mole fractions (see Section 2), thus we compare all of the systems on an equivalent basis. It is possible, however, that a bias in our influence functions contributes to the MBE in Figure 1, and yields incorrect rankings among these inversions.

In summary, we find the NEE of CO_2 in central and eastern North America by nearly all these inversion systems to be positively biased in summer and negatively biased in the other three seasons, with the degree of bias varying across the inversion system. Therefore, the magnitude of the seasonal cycle of NEE of CO_2 across central and eastern Temperate North America is likely to be underestimated across the models in the OCO-2 v9 MIP. The overall annual bias from these systems is not clear, since the seasonal flux biases change sign and will cancel out over the course of a year to a degree that is not clear from this analysis.

A number of broad patterns emerge when the MBE is evaluated for each ecoregion (Figure 2). 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 large MBEs are located in the Appalachian forests (ecoregion 5), central crops and forest (ecoregion 6), the corn belt (ecoregion 7), and the northern crops (ecoregion 8). More pronounced MBE levels in the positive and negative direction are found in the Baker and UT models, which may imply a smaller model-data-mismatch covariance given in the model than others. The OU model MBE most often diverges 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 when limiting the discussion to the IS and LNLG inversions.

During the summertime, we identify large positive biases in Appalachian forests (ecoregion 5), central crops and forests (ecoregion 6), the corn belt (ecoregion 7), and northern crops (ecoregion 8). The UT and Baker-mean models contain many of the peak positive biases across these ecoregions. The TM5-4DVAR model shows the smallest MBE

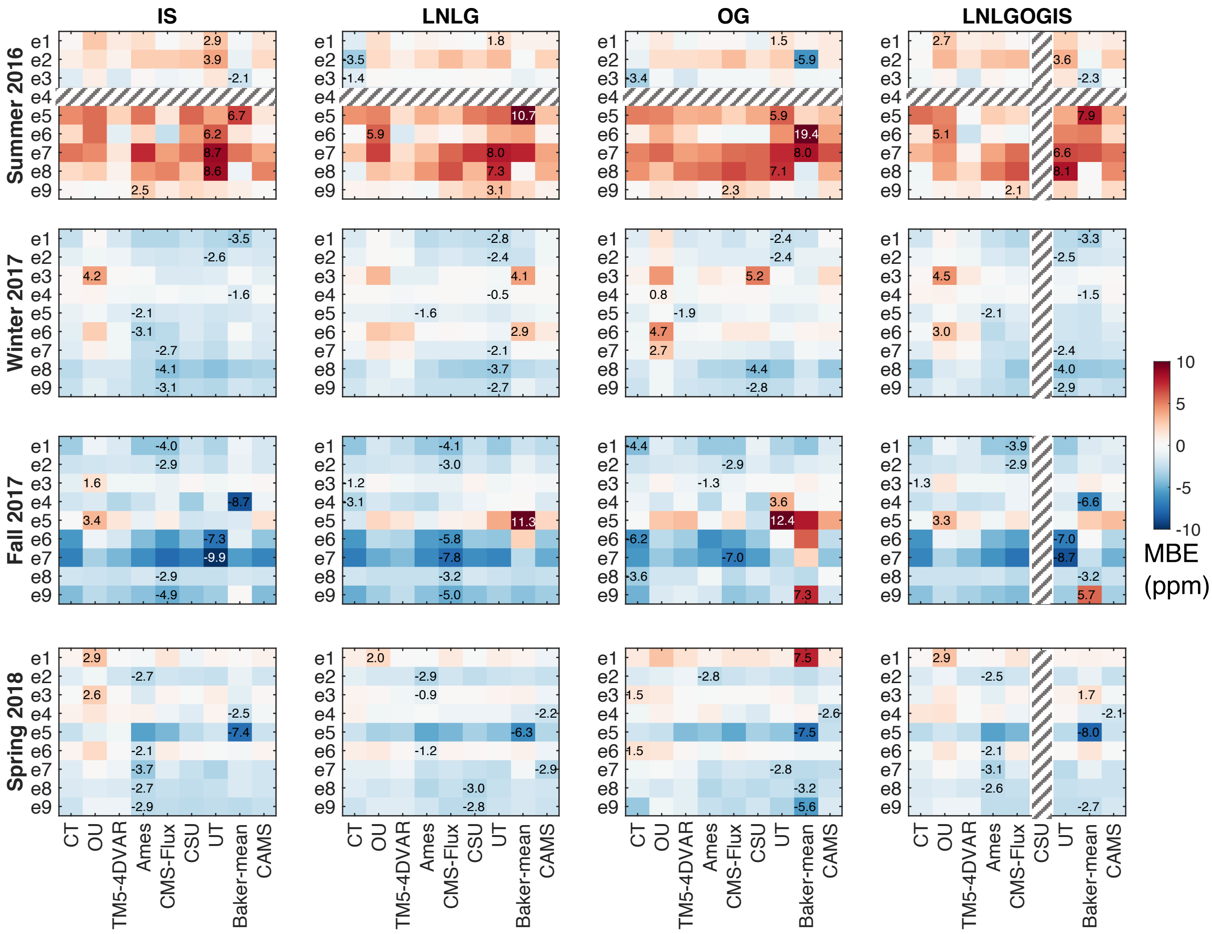

Figure 2. Mean Bias Error (MBE, ppm) for 9 different ecoregions in Central and Eastern Temperate North America. The largest magnitude of MBE for each ecoregion is written onto the cell. A warm color denotes a positive bias, and a cold color denotes a negative bias. The ecoregions are defined in Figure 2.Shaded areas denote no data.

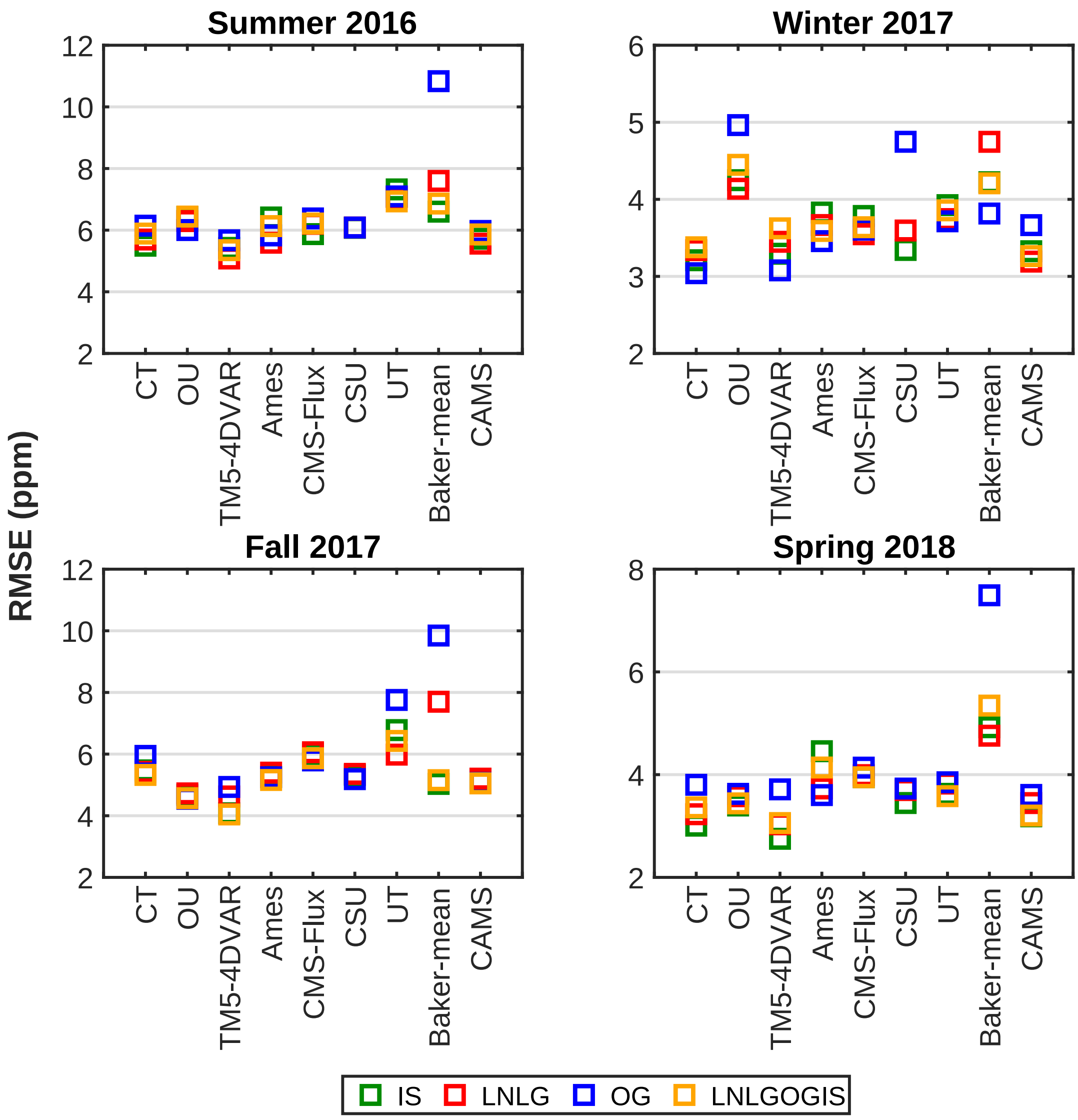

Figure 3. RMSE of the posterior biogenic CO_2 computed from all inverse estimates of NEE of CO_2 compared to the observed ABL CO_2 mole fractions from each of four seasonal ACT-America campaigns.

across all ecoregions. During the fall months, large negative MBE values are found in most ecoregions with the exception of the Appalachian forests (ecoregion 5) where the MBEs are positive. The Baker model again stands out in comparison to other inversion systems, with positive MBEs for many ecoregions when driven by OCO-2 data (i.e., LNLG and OG). Given that only moderate NEE of CO_2 is expected in the fall, the performance of the OCO-2 v9 MIP models during the fall months is relatively poor compared to other seasons.

Figure 3 and Figure 4 show the RMSE and MBE analysis, respectively, for four data experiments in OCO-2 v9 MIP including IS, LNLG, the OCO-2 v9 ocean glint ("OG") experiment, and the combination of IS, LNLG and OG experiment ("LNLGOGIS") (see details in section 2).

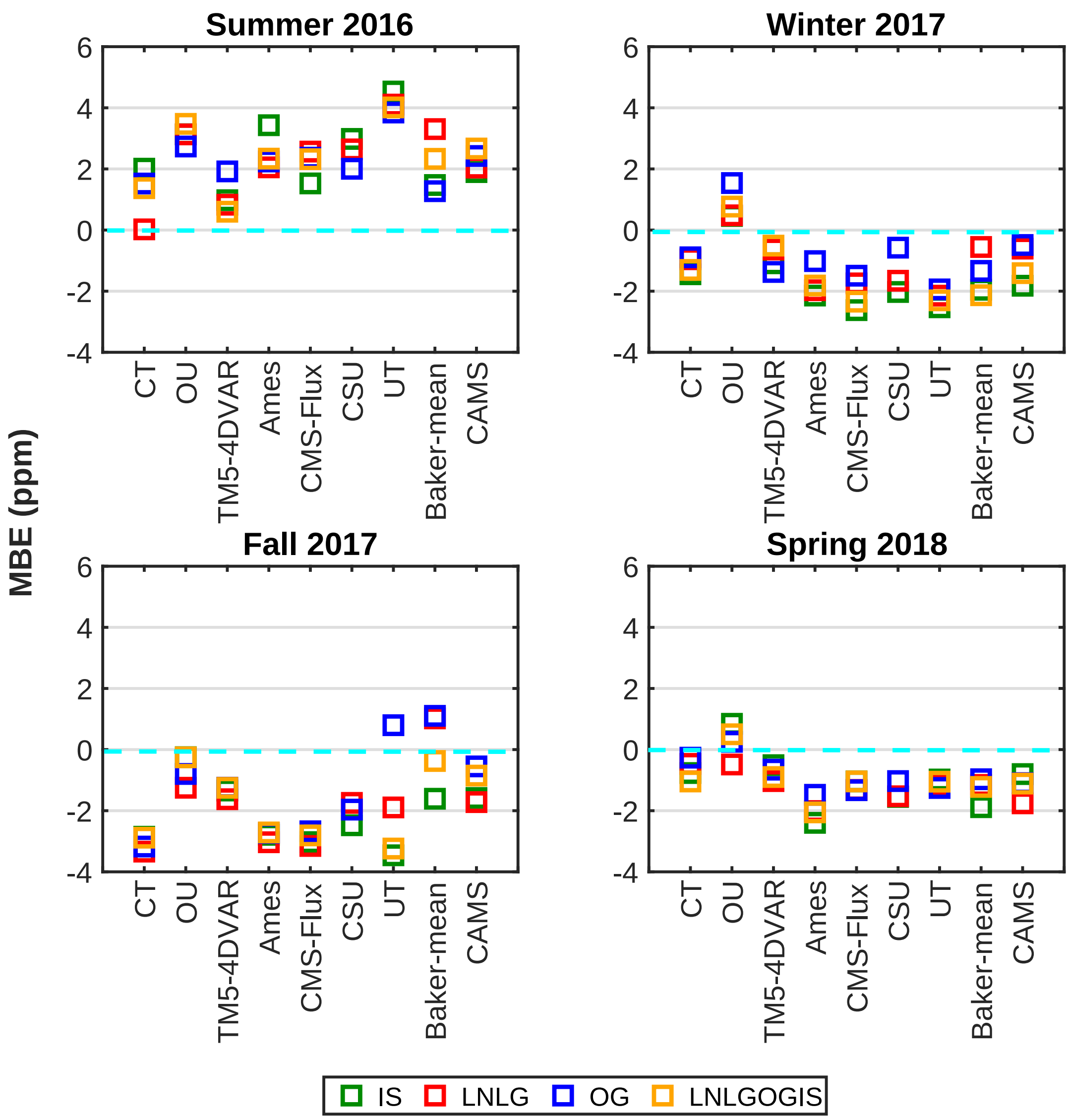

Figure 4. MBE of the posterior biogenic CO_2 computed from all inverse estimates of NEE of CO_2 compared to the observed ABL CO_2 mole fractions from each of four seasonal ACT-America campaigns.

The RMSE analysis (Figure 3) shows seasonal patterns likely related to flux magnitudes. 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. The variability of RMSE levels across different models is small during the spring months, and largest during the fall months. These findings are roughly consistent with larger NEE (Figure 1 and Figure S4-7), hence larger potential for model-data differences, in the more biologically active seasons.

Sensitivity of RMSE in the CO_2 NEE flux estimates to data sources varies, perhaps indicative of the construction of the 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 OCO-2 ocean glint v9 retrievals ("OG"). 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. In addition, RMSE analyses suggest that the OCO-2 v9 OG-based inversion is inferior to other experiments, yielding the highest RMSE across seasons and models.

The MBE analysis as a function of the observational data set shows similar patterns (Figure 4) 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. Unlike the RMSE analysis, the OG experiments here don't show the large discrepancies as compared to the other experiments. 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 superior data density. Patterns of MBE across models and regions have been discussed earlier in the paper.

4 Discussions and Conclusion

We implement a regional evaluation of net ecosystem exchange (NEE) of CO_2 flux products from nine current state-of-the-science global inversion systems in central and eastern temperate North America, using the largest carbon-centric, regional-scale aircraft mission (ACT-America) yet deployed anywhere on the earth. We estimate the seasonal performance of CO_2 NEE flux products of the OCO-2 v9 MIP across this portion of North America and expand the evaluation to the ecoregions within the domain.

The seasonal bias analysis shows that the inversion models' NEE estimates are positively biased in summer, and negatively biased in winter, fall, and spring across most flux products, suggesting that the seasonal magnitude of CO_2 NEE is underestimated in these global CO_2 inversion systems. The performance of the OCO-2 v9 land nadir/land glint data experiment is similar to the in situ experiment, an encouraging finding for regions of the world where the in situ observing network is sparse. The spatially resolved errors for the regional fluxes in the inversion models are not strongly dependent on the observational data sources for most of the models but a small number of the inversion systems display noticeably greater sensitivity to the data source. Large seasonal MBE values exist in the crop land and eastern forest regions.

The implication that most OCO-2 v9 MIP models underestimate the seasonal amplitude of NEE across central and eastern US ecosystems, regardless of data set, is striking. Similar results were found in two additional studies using ACT-America observations. Zhang et al. (2021) compared the posterior CO_2 4D fields from different inversion systems of OCO_2 v9 MIP to the ACT-America flight observations to understand the weather-driven atmospheric CO_2 differences, which does not separate the impacts of transport and flux errors. Zhang et al. (2021) found that most inversion systems in most seasons underestimated the difference in CO_2 between the ABL and the free troposphere, a result that is potentially consistent with systematically underestimated seasonal flux magnitudes. It is worth noting that the methods of Zhang et al. (2021) do not depend on a "third-party" atmospheric transport model to project mole fractions into flux space, as was done in this study. Feng et al. (2021) found a systematic underestimate of summer 2016 net uptake of CO_2 when comparing an ecosystem flux ensemble and CarbonTracker posterior fluxes to ACT-America and NOAA tall tower CO_2 observations. The results of Feng et al. (2021) use a WRF-Chem atmospheric ensemble to transport flux estimates, presenting a third and independent treatment of atmospheric transport yet yielding similar findings, albeit only for the summer season. Finally, Hu et al. (2019) used independent aircraft vertical profiles of CO_2 to evaluate CarbonTracker's CO_2 NEE inversion products and show similar seasonal-biases pattern in terms of simulating the ABL CO_2 mole fractions.

The impact of this apparent underestimate in the seasonal cycle of fluxes on annually integrated NEE of CO_2 of North America is not clear but deserves additional investigation. It is also possible that this seasonal bias could directly impact or is indicative of features of these inversions that could impact NEE estimates in other regions of the globe. The finding that the TM5-based inversions appear on average to have smaller seasonal biases than the GEOS-Chem-based inversions is also potentially consistent with the findings of Schuh et al. (2019). Schuh et al. (2019) suggested that TM5 mixes more vigorously in the vertical than GEOS-Chem. This could lead to TM5-based inversions requiring stronger NEE of CO_2 to match ABL CO_2 observations, since seasonal fluxes would be diluted within a larger atmospheric mixing volume. Schuh et al. (2019) showed that, globally, these differences in atmospheric mixing led to large differences in inverse estimates of annual NEE of CO_2. We suggest that continued understanding of the causes of the biases at sub-continental scales found in this study will enable increased confidence not just in regional, seasonal NEE, but in global, annual NEE estimates.

Acknowledgments

The ACT-America project is a NASA Earth Venture Suborbital 2 project funded by NASA's Earth Science Division. We 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. MJ acknowledges the internal funding from NASA's Earth Science Research and Analysis Program . We 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. All ACT-America in situ data used in the manuscript can be found at the ORNL DAAC (`https://daac.ornl.gov/actamerica`). The FLEXPART v10.4 model can be found online (`https://www.flexpart.eu/wiki/FpInstall`). The authors declare, to their knowledge, no conflicts of interest with the submission of this manuscript.

References

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., ... Tans, P. P. (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.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., ... Ziehn, T. (2020). Carbon–concentration and carbon–climate feedbacks in cmip6 models and their comparison to cmip5 models. *Biogeo-*

sciences, *17*(16), 4173–4222. Retrieved from https://bg.copernicus.org/articles/17/4173/2020/ doi: 10.5194/bg-17-4173-2020

Chevallier, F. (2021). Fluxes of carbon dioxide from managed ecosystems estimated by national inventories compared to atmospheric inverse modeling. *Geophysical Research Letters*, *48*(15), e2021GL093565. doi: 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. doi: 10.5194/acp-19-14233-2019

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., & Ramonet, M. (2010). Atmospheric inversions for estimating co2 fluxes: methods and perspectives. *Climatic Change*, *103*(1-2), 69 - 92. Retrieved from https://hal.inrae.fr/hal-02665340 doi: 10.1007/s10584-010-9909-3

Crisp, D., Miller, C. E., & DeCola, P. L. (2008). NASA Orbiting Carbon Observatory: measuring the column averaged carbon dioxide mole fraction from space. *Journal of Applied Remote Sensing*, *2*(1), 1 – 14. Retrieved from https://doi.org/10.1117/1.2898457 doi: 10.1117/1.2898457

Crowell, S., Baker, D., Schuh, A., Basu, S., Jacobson, A. R., Chevallier, F., ... Jones, D. B. A. (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.

Cui, Y. Y., Jacobson, A. R., Feng, S., Wesloh, D., Barkley, Z. R., Zhang, L., ... Davis, K. J. (2021). Evaluation of carbontracker's inverse estimates of north american net ecosystem exchange of co2 from different observing systems using act-america airborne observations. *Journal of Geophysical Research: Atmospheres*, *126*(12), e2020JD034406. doi: https://doi.org/10.1029/2020JD034406

Davis, K., Browell, E., Feng, S., Lauvaux, T., Obland, M., Pal, S., ... Williams, C. A. (2021). The atmospheric carbon and transport (act)-america mission. *Bull. Amer. Meteorol. Soc.*, 1–54. doi: https://doi.org/10.1175/BAMS-D-20-0300.1

Feng, S., Lauvaux, T., Williams, C. A., Davis, K. J., Zhou, Y., Baker, I., ... Wesloh, D. (2021). Joint co2 mole fraction and flux analysis confirms missing processes in casa terrestrial carbon uptake over north america. *Global Biogeochemical Cycles*, *35*(7), e2020GB006914. (e2020GB006914 2020GB006914) doi: https://doi.org/10.1029/2020GB006914

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., ... Yin, Y. (2019). Global atmospheric co_2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate. *Biogeosciences*, *16*(1), 117–134. Retrieved from https://bg.copernicus.org/articles/16/117/2019/ doi: 10.5194/bg-16-117-2019

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., ... Grainger, C. A. (2003). Toward constraining regional-scale fluxes of co2 with atmospheric observations over a continent: 1. observed spatial variability from airborne platforms. *Journal of Geophysical Research: Atmospheres*, *108*(D24). doi: https://doi.org/10.1029/2002JD003018

Hu, L., Andrews, A. E., Thoning, K. W., Sweeney, C., Miller, J. B., Michalak, A. M., ... van der Velde, I. R. (2019). Enhanced north american carbon uptake associated with el niño. *Science Advances*, *5*(6).

Karion, A., Callahan, W., Stock, M., Prinzivalli, S., Verhulst, K., Kim, J., ... Whetstone, J. (2020). Greenhouse gas observations from the northeast corridor tower network. *Earth Systems Science Data*.

Liu, J., Baskaran, L., Bowman, K., Schimel, D., Bloom, A. A., Parazoo, N. C., ... Wofsy, S. (2021). Carbon monitoring system flux net biosphere exchange

2020 (cms-flux nbe 2020). *Earth System Science Data*, *13*(2), 299–330. doi: 10.5194/essd-13-299-2021

Liu, J., & Bowman, K. (2016). A method for independent validation of surface fluxes from atmospheric inversion: Application to co2. *Geophysical Research Letters*, *43*(7), 3502-3508. doi: 10.1002/2016gl067828

Millar, R., Fuglestvedt, J., Friedlingstein, P., Rogelj, J., Grubb, M., Matthews, H., ... Allen, M. (2017). Emission budgets and pathways consistent with limiting warming to 1.5 c. *Nature Geoscience*, *10*, 741-747.

Miller, C. E., Griffith, P. C., Goetz, S. J., Hoy, E. E., Pinto, N., McCubbin, I. B., ... Margolis, H. A. (2019). An overview of ABoVE airborne campaign data acquisitions and science opportunities. *Environmental Research Letters*, *14*(8), 080201.

Miller, J. B., Lehman, S. J., Verhulst, K. R., Miller, C. E., Duren, R. M., Yadav, V., ... Sloop, C. D. (2020). Large and seasonally varying biospheric co2 fluxes in the los angeles megacity revealed by atmospheric radiocarbon. *Proceedings of the National Academy of Sciences*, *117*(43), 26681–26687. doi: 10.1073/pnas.2005253117

Pal, S., K. J. Davis, K. J., Pauly, R. M., McGill, M. J., Campbell, L., & Hoffman, K. (2020). A brief description of the cloud physics lidar-derived atmospheric boundary layer top height data sets obtained using wavelet transform algorithm. *ORNL DAACm, Oak Ridge, Tennessee, USA.*.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., ... Hayes, D. (2011). A large and persistent carbon sink in the world's forests. *Science*, *333*(6045), 988–993. doi: 10.1126/science.1201609

Peiro, H., Crowell, S., Schuh, A., Baker, D. F., O'Dell, C., Jacobson, A. R., ... Baker, I. (2021). Four years of global carbon cycle observed from oco-2 version 9 and *in situ* data, and comparison to oco-2 v7. *Atmospheric Chemistry and Physics Discussions*, *2021*, 1–50. Retrieved from `https://acp.copernicus.org/preprints/acp-2021-373/` doi: 10.5194/acp-2021-373

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., ... Zhang, X. (2013). Global atmospheric carbon budget: results from an ensemble of atmospheric co_2 inversions. *Biogeosciences*, *10*(10), 6699–6720. doi: 10.5194/bg-10-6699-2013

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., Cassiani, M., Eckhardt, S., ... Stohl, A. (2019). The lagrangian particle dispersion model flexpart version 10.4. *Geoscientific Model Development*, *12*(12), 4955–4997. doi: 10.5194/gmd-12-4955-2019

Rayner, P. J., Michalak, A. M., & Chevallier, F. (2019). Fundamentals of data assimilation applied to biogeochemistry. *Atmospheric Chemistry and Physics*, *19*(22), 13911–13932. doi: 10.5194/acp-19-13911-2019

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., ... Palmer, P. I. (2019). Quantifying the impact of atmospheric transport uncertainty on co2 surface flux estimates. *Global Biogeochemical Cycles*, *33*(4), 484-500. doi: https://doi.org/10.1029/2018GB006086

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., ... Tans, P. P. (2015). Seasonal climatology of co2 across north america from aircraft measurements in the noaa/esrl global greenhouse gas reference network. *Journal of Geophysical Research: Atmospheres*, *120*(10), 5155-5190. doi: https://doi.org/10.1002/2014JD022591

Tan, Z., Liu, S., Sohl, T. L., Wu, Y., & Young, C. J. (2015). Ecosystem carbon stocks and sequestration potential of federal lands across the conterminous united states. *Proceedings of the National Academy of Sciences*, *112*(41), 12723–12728. doi: 10.1073/pnas.1512542112

Wang, J., Feng, L., Palmer, Y., Paul I.and Liu, Fang, S., Bösch, H., O'Dell,

C. W., . . . Xia, C. (2020). Large chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature*, *586*(7831), 720-723. doi: 10.1038/s41586-020-2849-9

Wei, Y., Shrestha, R., Pal, S., Gerken, T., McNelis, J., Singh, D., . . . Davis., K. (2021). The atmospheric carbon and transport (act) – america datasets: Description, management, and delivery. *Earth and Space Science*. doi: https://www.essoar.org/doi/pdf/10.1002/essoar.10505692.1.

Zhang, L., Davis, K. J., Schuh, A. E., Jacobson, A. R., Pal, S., Cui, Y. Y., . . . Basu, S. (2021). Multi-season evaluation of co2 weather in oco-2 mip models. *Earth and Space Science Open Archive*, 29. doi: 10.1002/essoar.10507526.1

Supporting Information for "Seasonal strength of terrestrial net ecosystem CO_2 exchange from North America is underestimated in global inverse modeling"

Yu Yan Cui[1], Li Zhang[1], Andrew R. Jacobson[2,3], Matthew S. Johnson[4], Sajeev Philip[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]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

[3]NOAA Earth System Research Laboratory, Global Monitoring Laboratory, Boulder, CO, USA

[4]Earth Science Division, NASA Ames Research Center, Moffett Field, CA, USA

[5]Universities Space Research Association, NASA Ames Research Center, Mountain View, CA, USA

[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]University of Oklahoma, School of Meteorology, Norman, OK, USA

[10]Department of Physics, University of Toronto, Toronto, ON, Canada

[11]NASA GSFC GMAO / University of Maryland, Goddard Space Flight Center, Greenbelt, MD, USA

[12]Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA, USA

Contents of this file

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

Text S1.

The background CO_2 mole fractions (C_{bkg}) for each receptor are determined by combining the sensitivity of each receptor to the initial condition (m, prior to the backward 10 days) and the OCO-2 v9 MIP global optimized CO_2 mole fraction fields (C_{CO_2}) (Equation S1).An example of the background determination is shown in Figure S1.

$$C_{bkg} = m \cdot C_{CO_2} \tag{1}$$

where m is the spatially and temporally resolved sensitivity field of the receptors to the initial conditions (dimension: n $\times$ i $\times$ j $\times$ z, n denotes the receptors, i, j, and z are latitude, longitude, and altitude, respectively), C_{CO_2} is the corresponding inversion-optimized CO_2 mole fraction fields (dimension: i $\times$ j $\times$ z), and C_{bkg} is the determined background values for each receptor (dimension: n $\times$ 1).

We define 12 ecoregions in the study (Figure S2) and calculated the influence functions with these ecoregions. For each receptor, the ecoregion associated with the largest contribution to the influence functions is tagged as the representative region (Figure S2).

We calculate the seasonal NEE flux budget (PgC/yr) for the shaded areas (Figure S3) and analyze the relationships between the regional flux strength and the estimates of Mean Bias Error based on the ACT-America aircraft campaigns.

The maps of averaged CO_2 NEE during the ACT-America campaign months from the inversion products are shown in Figure S4, Figure S5,Figure S6, and Figure S7.

As mentioned in the main text, a suite of gridded global CO_2 NEE products at 3-hourly resolution from nine members of OCO-2 v9 MIP (Table S1) was created for the four ACT-America Campaign periods (summer 2016, winter 2017, fall 2017, and spring 2018).

The global atmospheric CO_2 inversion models are driven by different prior flux components including Fossil fuel, NEE, fire and ocean fluxes. All models used the same fossil fuel flux products from ODIAC 2018 version (`https://gmao.gsfc.nasa.gov/gmaoftp/sourish/ODIAC/2018/distrib/`), and the prior flux from NEE, fire, and ocean components are listed in Table S2.

Table S1. Basic information of the nine global inversion systems evaluated in the study.

Inversion system	Transport model	Met driver	Inversion method	Flux spatial resolution	Flux temporal resolution
CT	TM5	ERA-Interim	EnKF	1x1	3-hourly
OU	TM5	ERA-Interim	4DVar	1x1	3-hourly
TM5-4DVAR	TM5	ERA-Interim	4DVar	2x3	3-hourly
Ames	GEOS-Chem	MERRA-2	4DVar	4x5	3-hourly
CMS-Flux	GEOS-Chem	MERRA-2	4Dvar	4x5	3-hourly
CSU	GEOS-Chem	GEOS-FP	Beyesian synthesis	1x1	3-hourly
UT	GEOS-Chem	GEOS-FP	4Dvar	4x5	3-hourly
Baker-mean	PCTM	MERRA-2	4Dvar	2x2.5	3-hourly
CAMS	LMDZ3	ERA5	Variational	1.875 x 3.75	3-hourly

Table S2. Prior inventories of CO_2 flux components

	NEE	Ocean	Fire
CT	CT2019 CASA-GFED4	CT2019 OI	CT2019 w4 (based on GFED4)
OU	CASA-GFED3	Takahashi	GFED3
TM5-4DVAR	SiB4	Lofi	GFED4
Ames	CT2019 CASA-GFED4	CT2019 OI	CT2019 w4 (based on GFED4)
CMS-Flux	CARDAMOME	ECCO-Darwin	GFED3
CSU	SIB4/MERRA	CT2015 OI	GFED4-HEMCO
UT	CASA-GFED4	Takahashi	GFED4
Baker-Mean	CASA-GFED4	Takahashi	GFED3
CAMS	ORCHIDEE	CMEMS	GFED4

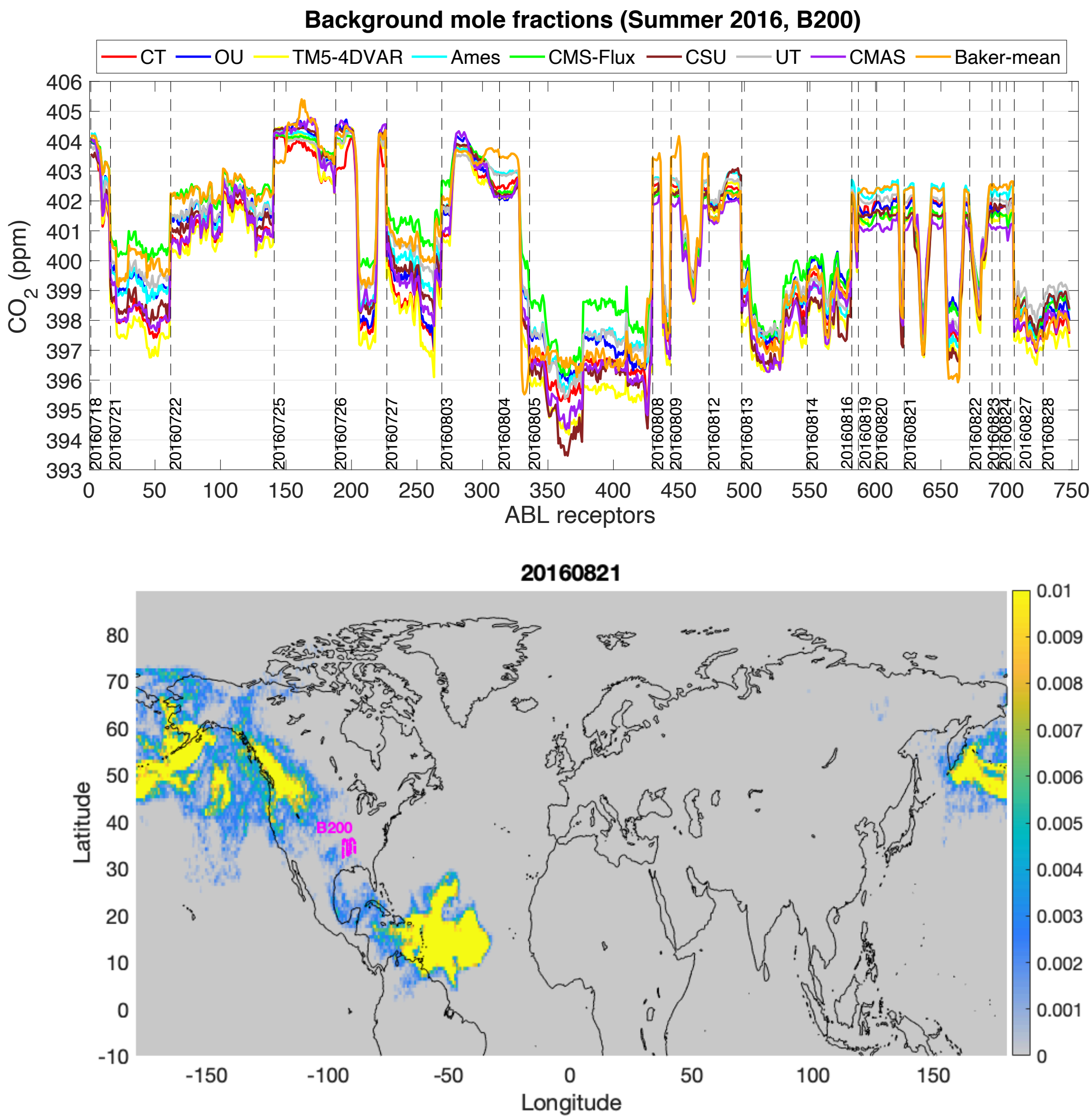

Figure S1. The upper panel shows the background CO2 mole fractions for each receptor from the all B200 flights during the ACT summer 2016 campaign for each OCO-2 v9 MIP model that is associated with the IS experiment. The lower panel shows the integrated sensitivity of all receptors from one single B200 flight (pink line is the flight track) to the initial condition (backward 10 days) across all vertical layers. Background CO_2 (C_{bkg}) for all ACT airborne ABL observations were computed in this fashion.

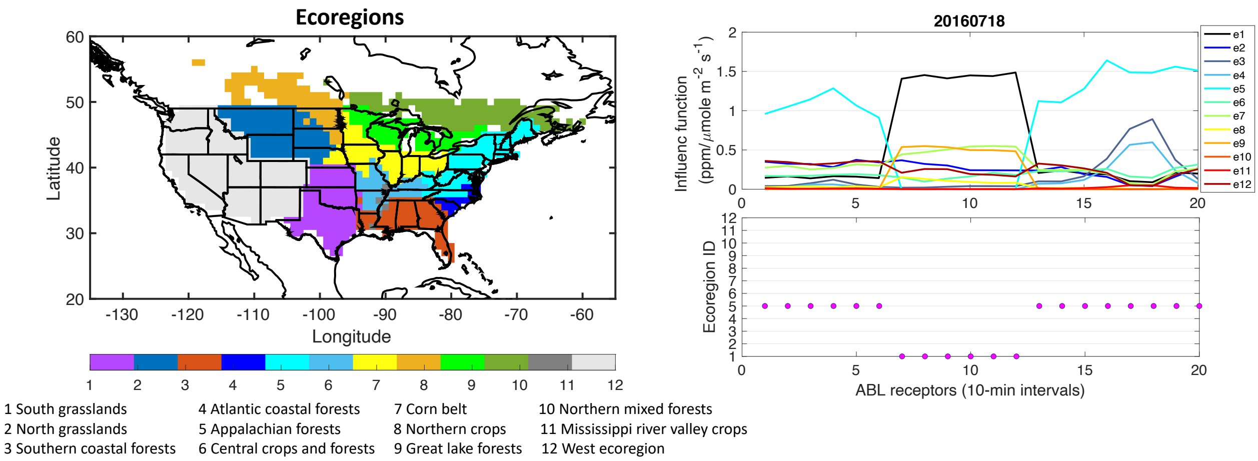

Figure S2. The left panel displays the spatial patterns of ecoregions in Temperate North America defined in the study. The right panel presents an example of this ecoregion tagging process for each receptor in the ABL of one ACT flight.

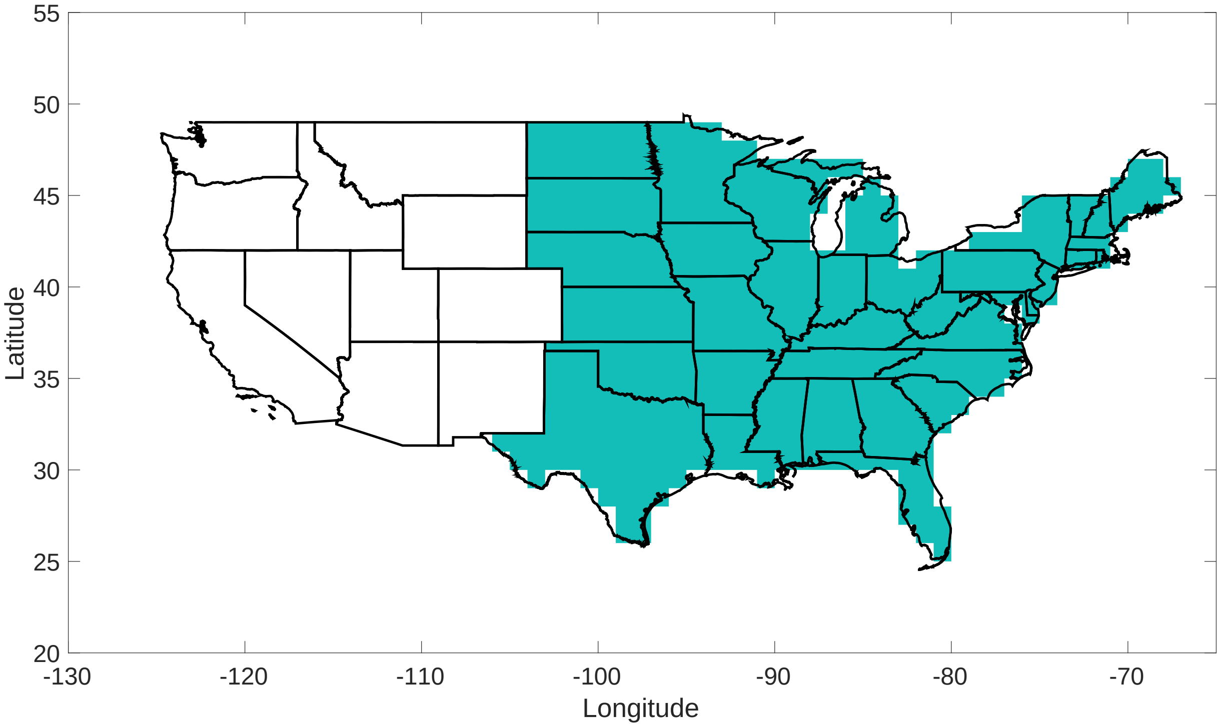

Figure S3. The seasonal NEE flux strength from the shaded areas are calculated for Figure 3.

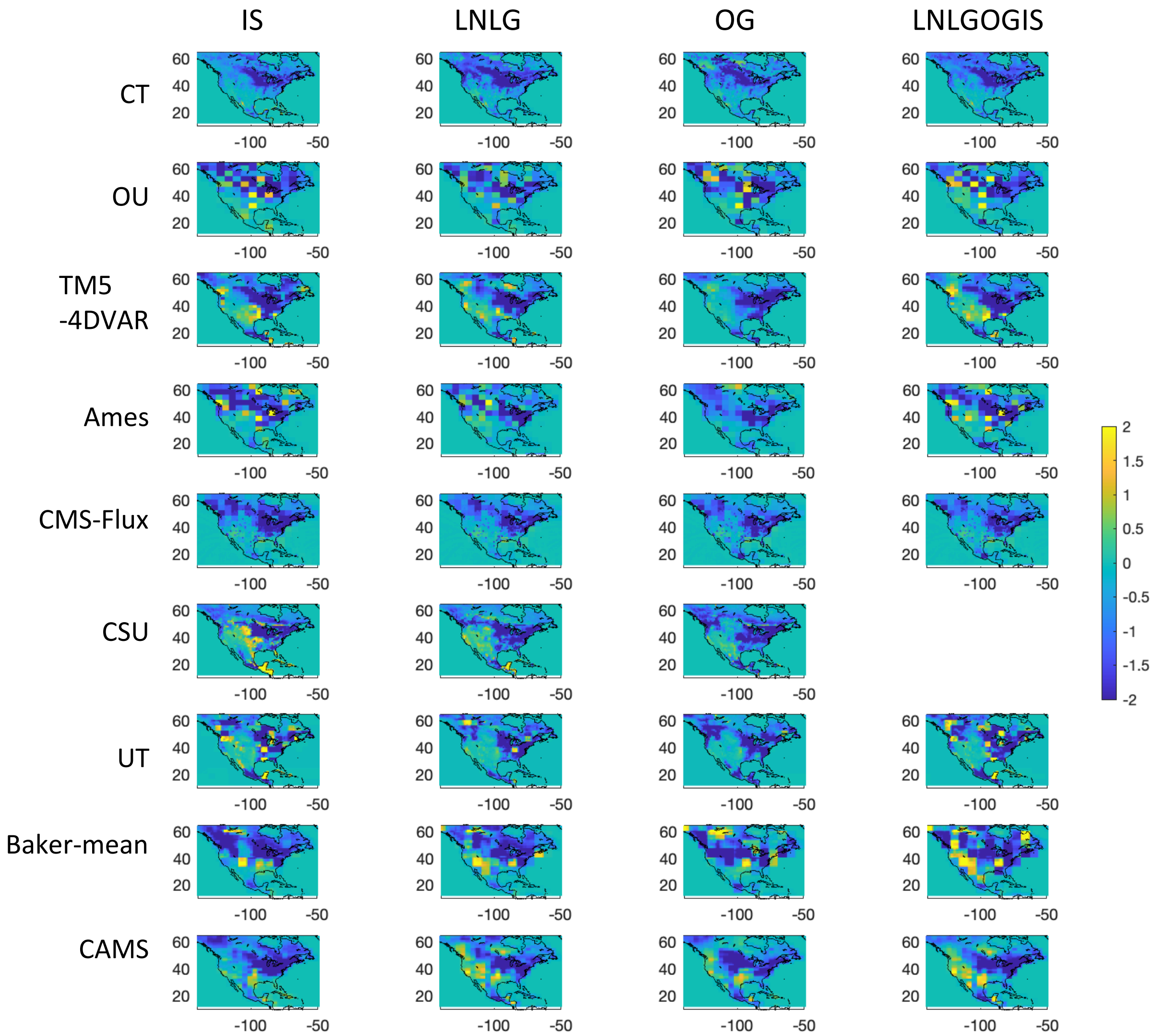

Figure S4. Averaged CO2 NEE fluxes (unit:µmol m-2 s-1) during July-August 2016.

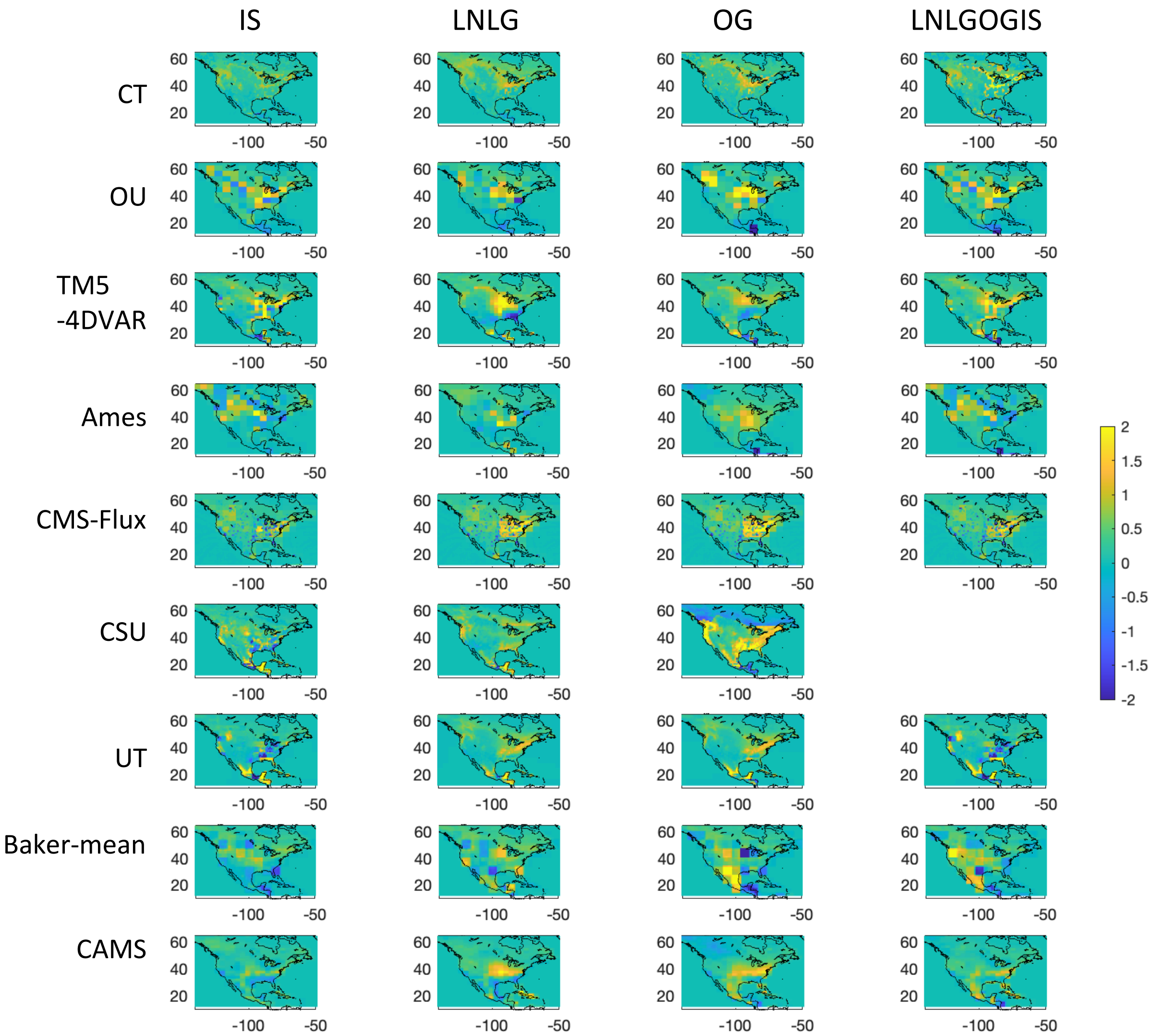

Figure S5. Averaged CO2 NEE fluxes (unit:μmol m-2 s-1) during February-March 2017.

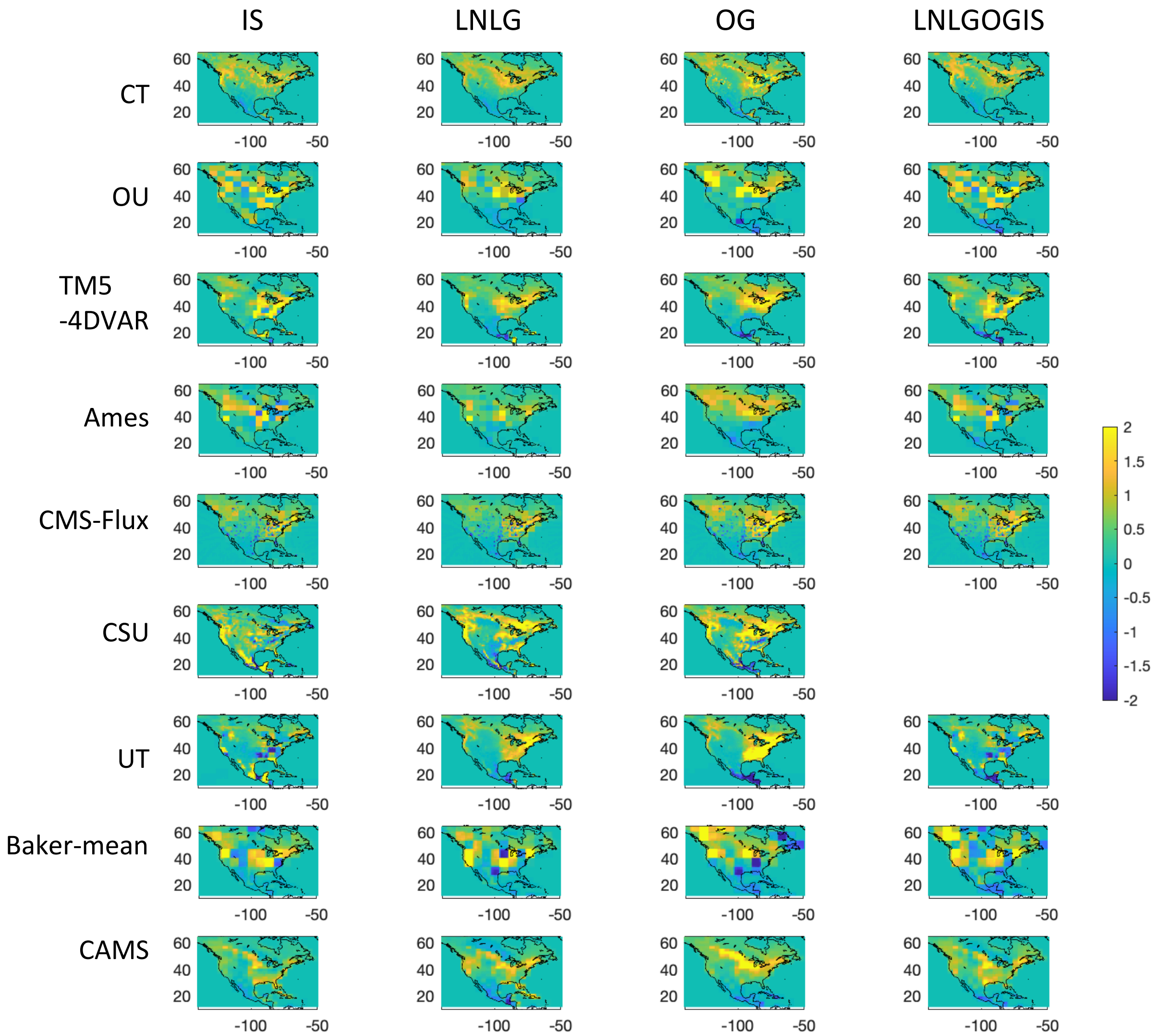

Figure S6. Averaged CO2 NEE fluxes (unit:µmol m-2 s-1) during October-November 2017.

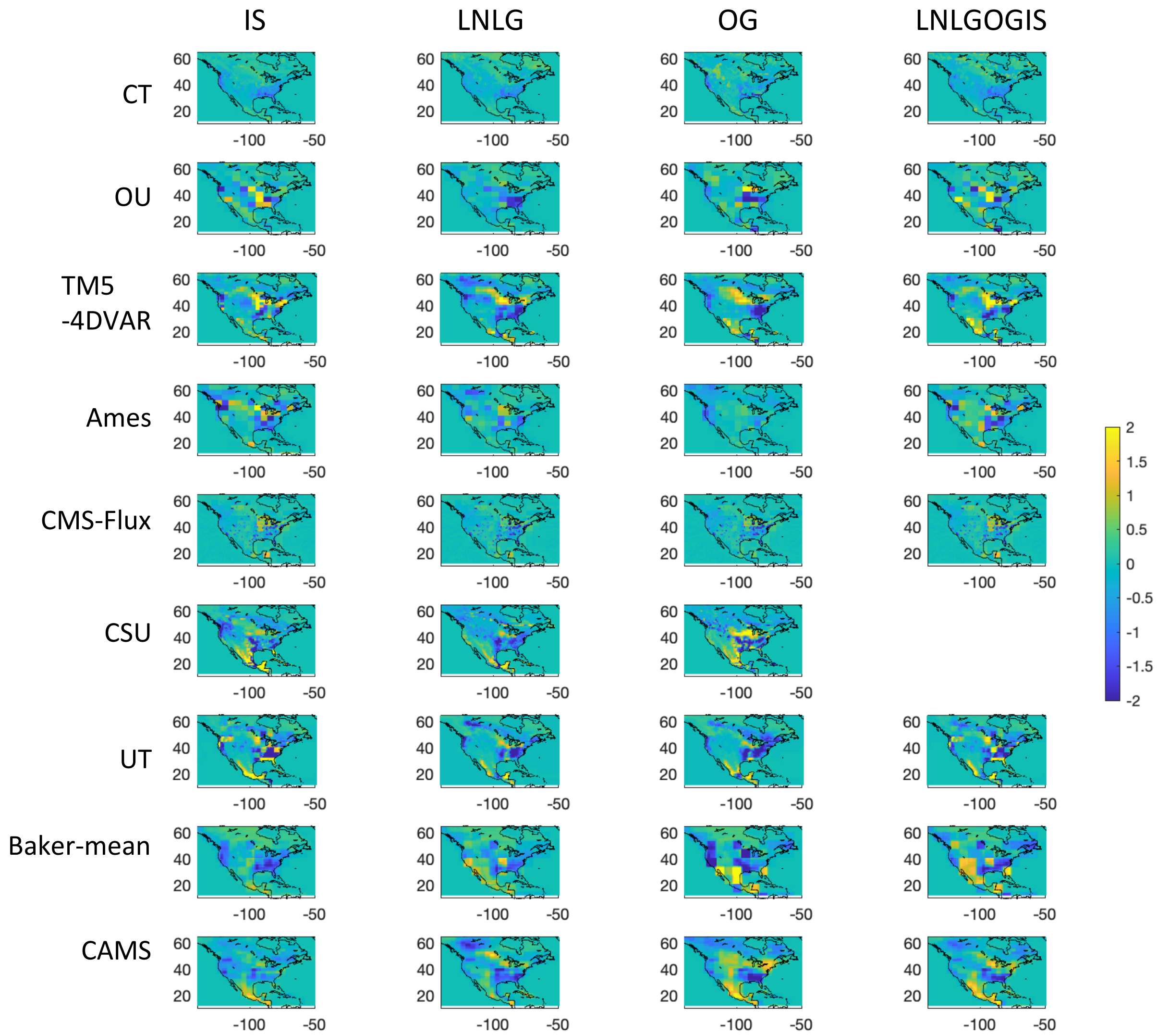

Figure S7. Averaged CO2 NEE fluxes (unit:µmol m-2 s-1) during April-May 2018