



HAL
open science

Joint CO₂ Mole Fraction and Flux Analysis Confirms Missing Processes in CASA Terrestrial Carbon Uptake Over North America

Sha Feng, Thomas Lauvaux, Christopher A. Williams, Kenneth J. Davis, Yu Zhou, Ian T. Baker, Zachary R. Barkley, Daniel Wesloh

► **To cite this version:**

Sha Feng, Thomas Lauvaux, Christopher A. Williams, Kenneth J. Davis, Yu Zhou, et al.. Joint CO₂ Mole Fraction and Flux Analysis Confirms Missing Processes in CASA Terrestrial Carbon Uptake Over North America. *Global Biogeochemical Cycles*, 2021, 35 (7), 10.1029/2020GB006914 . hal-03329882

HAL Id: hal-03329882

<https://hal.science/hal-03329882>

Submitted on 16 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Global Biogeochemical Cycles

RESEARCH ARTICLE

10.1029/2020GB006914

Special Section:

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

Key Points:

- In summer, uncertainty in biogenic fluxes is the largest source of uncertainty in modeled atmospheric CO₂
- Atmospheric CO₂ mole fraction and eddy-flux measurements confirm a sink over North America not represented in the Carnegie-Ames-Stanford-Approach (CASA) terrestrial model
- Eddy-flux measurements suggest CASA summer daytime uptake is too weak and nighttime respiration is too strong

Supporting Information:

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

Correspondence to:

S. Feng,
sfeng@pnnl.gov

Citation:

Feng, S., Lauvaux, T., Williams, C. A., Davis, K. J., Zhou, Y., Baker, I., et al. (2021). Joint CO₂ mole fraction and flux analysis confirms missing processes in CASA terrestrial carbon uptake over North America. *Global Biogeochemical Cycles*, 35, e2020GB006914. <https://doi.org/10.1029/2020GB006914>

Received 9 DEC 2020

Accepted 11 JUN 2021

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Joint CO₂ Mole Fraction and Flux Analysis Confirms Missing Processes in CASA Terrestrial Carbon Uptake Over North America

Sha Feng^{1,2} , Thomas Lauvaux^{1,3} , Christopher A. Williams⁴ , Kenneth J. Davis^{1,5} , Yu Zhou⁴ , Ian Baker⁶, Zachary R. Barkley¹ , and Daniel Wesloh¹ 

¹Department of Meteorology and Atmospheric Science, The Pennsylvania State University, University Park, PA, USA,

²Atmospheric Sciences and Global Change division, Pacific Northwest National Laboratory, Richland, WA, USA,

³Laboratoire des Sciences du Climat et de l'Environnement, CEA, CNRS, UVSQ/IPSL, Université Paris-Saclay, Orme des Merisiers, Gif-sur-Yvette, France, ⁴Graduate School of Geography, Clark University, Worcester, MA, USA, ⁵Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA, USA, ⁶Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA

Abstract Terrestrial biosphere models (TBMs) play a key role in the detection and attribution of carbon cycle processes at local to global scales and in projections of the coupled carbon-climate system. TBM evaluation commonly involves direct comparison to eddy-covariance flux measurements. We use atmospheric CO₂ mole fraction ([CO₂]) measured in situ from aircraft and tower, in addition to flux-measurements from summer 2016 to evaluate the Carnegie-Ames-Stanford-Approach (CASA) TBM. WRF-Chem is used to simulate [CO₂] using biogenic CO₂ fluxes from a CASA parameter-based ensemble and CarbonTracker version 2017 (CT2017) in addition to transport and CO₂ boundary condition ensembles. The resulting “super ensemble” of modeled [CO₂] demonstrates that the biosphere introduces the majority of uncertainty to the simulations. Both aircraft and tower [CO₂] data show that the CASA ensemble net ecosystem exchange (NEE) of CO₂ is biased high (NEE too positive) and identify the maximum light use efficiency E_{\max} a key parameter that drives the spread of the CASA ensemble in summer 2016. These findings are verified with flux-measurements. The direct comparison of the CASA flux ensemble with flux-measurements confirms missing sink processes in CASA. Separating the daytime and nighttime flux, we discover that the underestimated net uptake results from missing sink processes that result in overestimation of respiration. NEE biases are smaller in the CT2017 posterior biogenic fluxes, which assimilate observed [CO₂]. Flux tower analyses reveal an unrealistic overestimation of nighttime respiration in CT2017 which we attribute to limited flexibility in the inversion strategy.

1. Introduction

The Paris agreement calls for “a balance between anthropogenic emissions by sources and removals by sinks of greenhouse gases in the second half of this century” to limit the increase in global average temperature below 2°C above preindustrial levels (UNFCCC, 2015). Monitoring and tracking this commitment require accurate quantification of terrestrial carbon exchange with the atmosphere. The rapid increase of carbon dioxide (CO₂) has been partially offset by natural biogeochemical processes, including uptake by terrestrial and oceanic ecosystems (Friedlingstein et al., 2019). Understanding terrestrial carbon exchange is crucial for assessing biosphere-atmosphere interactions, for diagnosing terrestrial ecosystem contributions to the global coupled carbon-climate system and, ultimately, for reducing the uncertainty in climate projections (Friedlingstein et al., 2014).

Though the net global flux of CO₂ to the atmosphere is well-constrained (Tans & Conway, 2005; Tans et al., 1990), regional-to-continental biogenic CO₂ fluxes are not well characterized in current carbon estimation approaches (Crowell et al., 2019). Two approaches are commonly taken to quantify biogenic CO₂ fluxes at the continental scale. “Top-down” approaches use an optimization process by which atmospheric CO₂ mole fraction ([CO₂]) measurements combined with an atmospheric transport model are used to constrain the a priori estimation of the spatial and temporal distribution of biogenic CO₂ fluxes (Enting et al., 1995). Here we use “[]” to represent atmospheric mole fractions and distinguish it from flux space, and this convention is used hereafter. Determining and reducing the uncertainty associated with transport

and with flux priors have been the focus of the atmospheric inversion community (Baker et al., 2006; Gurney et al., 2002; Philip et al., 2019; Schuh et al., 2019). Flux priors are commonly obtained from the simulated net ecosystem exchange (NEE) from terrestrial biosphere models (TBMs).

TBMs (a “bottom up” approach) simulate surface CO₂ fluxes from site level to global scale as they integrate ecological and meteorological drivers (Fung et al., 1987). Flux priors in atmospheric inversions are commonly obtained from the NEE simulated by TBMs. TBMs also simulate the land carbon component in Earth system models used for climate projections. Joint assimilation systems have also been developed to benefit from both ecosystem measurements and atmospheric mole fractions (Kaminski et al., 2002). The optimization procedure relies on an adjoint model of the biogeochemical processes such as the Biosphere Energy-Transfer Hydrology model included in the Carbon Cycle Data Assimilation System at large scales (Knorr, 2000). Such a system was enhanced to assimilate satellite ecosystem products available over the globe (Kaminski et al., 2012) or eddy-flux and mole fraction measurements over continents (Koffi et al., 2013). These approaches optimize a number of model parameters in part of the underlying TBMs, and have been applied using remote [CO₂] from towers or satellites (e.g., Scholze et al., 2019). On sub-continental scales, atmospheric [CO₂] require fine-resolution models to simulate the complex atmospheric dynamics (e.g., Feng, Lauvaux, Keller, et al., 2019) and dense tower networks (Andrews et al., 2014) combined with eddy-flux tower networks (e.g., AmeriFlux) covering a wide array of ecosystems and climatic zones for evaluations

TBMs have been shown to vary widely in their projections of terrestrial CO₂ sink strengths, not only in magnitude but also even in sign (Baker et al., 2006; Gurney et al., 2002). Huntzinger et al. (2011) evaluated flux variability from four TBMs over North America and the potential impact on the inversion results. They found that the diurnal variability in surface fluxes within the near field of tower [CO₂] observations appear to have a significant impact on the high-frequency variations in the atmospheric data, and, thus, the inversion needs to adjust the temporal (and spatial) variability of the prior fluxes. Feng, Lauvaux, Keller, et al. (2019) compared the modeled [CO₂] errors attributed to biogenic CO₂ fluxes, fossil fuel emissions, atmospheric transport, and large-scale boundary inflow from daily to annual timescales and discovered that the biogenic CO₂ fluxes dominate the model errors across timescales, implying that [CO₂] observations hold promise for evaluating and improving TBMs. Global inversions usually optimize CO₂ fluxes on monthly to annual timescales at the grid scale in the magnitude of a few degrees by a few degrees. Higher spatiotemporal resolutions in the prior are often not altered. As the resolution of inversion system improves and the need for accurate flux data at high resolution increase, evaluating the higher-resolution spatial and temporal structure of TBMs is increasingly important. TBMs with more accurate, high-resolution surface biogenic CO₂ fluxes will improve our diagnosis, attribution, and projection of terrestrial carbon dynamics, and improve “top-down” analysis systems that rely on TBMs for a priori surface flux estimation.

Several multiple-model intercomparison projects have been conducted to characterize or synthesize current understanding of land-atmosphere carbon exchange and inform the uncertainty or confidence surrounding projections of future exchange and feedbacks with the climate system, such as the Multi-Scale Synthesis and Terrestrial Model Intercomparison Project (Huntzinger et al., 2013) Trends in Net Land-Atmosphere Carbon Exchange (TRENDY1; Sitch et al., 2008), the regionally focused Large Scale Biosphere Atmosphere-Data Model Intercomparison Project (LBAMIP2; Gonçalves et al., 2013), and the International Land-Atmosphere Benchmarking Project (ILAMB; Luo et al., 2012). These model intercomparison projects were built upon the protocols that specify standard model inputs, simulations and simulation setup procedures. These model intercomparison exercises have been used to explore the uncertainty in model simulations that arises from internal variability, boundary conditions, and parameter values for structural uncertainty from different model fluctuations (Schwalm et al., 2015). However, due to the complexity of TBMs, it is challenging to trace errors in individual models to misrepresentation of specific ecological processes or inappropriate model parameters through these model intercomparison projects. This merits deep exploration of just one modeling framework but with a perturbed parameter ensemble.

Recently, Zhou, Williams, Lauvaux, Davis, et al. (2020) introduced an ensemble of biogenic CO₂ fluxes simulated by the Carnegie-Ames-Stanford-Approach (CASA) biosphere model for North America at the resolutions of 5 km for North America and ~500 m for the US CONUS region from 2003 to 2019 by perturbing three model parameters—maximum light use efficiency E_{\max} , optimal temperature of photosynthesis

T_{opt} , and temperature response of respiration Q_{10} —on the basis of the CASA biome types. Those parameters were chosen as a result of a sensitivity test of simulated biogenic CO_2 fluxes to a series of CASA parameters. Furthermore, the range of E_{max} values was determined at the ecosystem level by comparison with eddy-covariance measurements of net CO_2 flux. Zhou, Williams, Lauvaux, Davis, et al. (2020) illustrated that the pruned L2 ensemble has good agreement with the flux data and outperforms many other TBMs at diurnal and annual scales, while overestimating model errors as represented by comparison with flux measurements.

Atmospheric $[\text{CO}_2]$ measurements provide another opportunity for evaluating modeled biogenic CO_2 fluxes, as evidenced by the fact that continental $[\text{CO}_2]$ gradients are mainly attributed to the biosphere (e.g., Feng, Lauvaux, Davies, et al., 2019; Feng, Lauvaux, Keller, et al., 2019). The major differences between mole fraction and flux measurements fall in the size of the surface influence area, or footprint, that influences a given measurement, and also the upwind memory of the samples. The size of a flux tower footprint is only about 1 km (e.g., McCaughey et al., 2006), and the measurements carry nearly instantaneous information of surface fluxes. Owing to these two factors, the flux measurements do not directly represent regional to continental fluxes, as the local fluxes captured may not be representative of broader scale patterns. One objective of this study is to explore the coherence between CO_2 flux and mole fractions with respect to modeled biogenic CO_2 flux evaluations. The CO_2 flux measurements selected for this study are from the AmeriFlux network (<https://ameriflux.lbl.gov/>), which has more than 150 active flux sites sampling a wide range of sites from the Amazonian rainforests to the North Slope of Alaska.

An in situ $[\text{CO}_2]$ tower typically has a footprint of hundreds of kilometers (e.g., Gloor et al., 2001; Sweeney et al., 2015) and carries the integrated atmospheric $[\text{CO}_2]$ signals from day and night. Aircraft measurements have even larger footprints with broad spatial sampling. TBM-modeled CO_2 fluxes can be evaluated against $[\text{CO}_2]$ measurements by means of Lagrangian and forward Eulerian transport modeling. Both are subject to transport model errors (e.g., Pillai et al., 2012). In Lagrangian modeling, the modeled biogenic CO_2 mole fractions ($[\text{CO}_2\text{bio}]$ hereafter) can be directly calculated by convolving biogenic CO_2 fluxes with the influencing areas of the $[\text{CO}_2]$ measurements that are simulated by an Eulerian transport model (Uliasz et al., 1994). The difficulty of this approach is to define the observed $[\text{CO}_2\text{bio}]$ due to mixed signals in the $[\text{CO}_2]$ measurements (e.g., Ogle et al., 2015). In Eulerian transport modeling, the CO_2 transport is treated as a passive tracer (Sarrat et al., 2007). The modeled total $[\text{CO}_2]$ is the sum of biogenic, fossil fuel, oceanic, fire CO_2 components in conjunction with boundary conditions as described in Feng, Lauvaux, Davies, et al. (2019) and Feng, Lauvaux, Keller, et al. (2019). The modeled error therefore can be from any or multiple components in addition to model transport (Feng, Lauvaux, Davies, et al., 2019; Feng, Lauvaux, Keller, et al., 2019). Note that these model error sources are also of concern in the “top-down” estimation. Another objective of this work is to explore to what degree the atmospheric $[\text{CO}_2]$ data can be used to evaluate TBMs. Here we use CASA (Potter et al., 1993) for demonstration and adopt the Eulerian transport model WRF-Chem (Grell et al., 2005; Skamarock et al., 2008) to serve this objective.

In this study, we employ both ground-based and aircraft in situ $[\text{CO}_2]$ data for the model evaluations. The NASA-funded Atmospheric Carbon and Transport (ACT)-America project was designed to improve the CO_2 and methane (CH_4) flux estimates by reducing transport and flux uncertainties. Two aircraft measured atmospheric CO_2 , CH_4 , and other gas species over Mid-Atlantic, Mid-West, and Southern Gulf regions in fair and frontal weather regimes. Typical flights encompassed 4–6 h of midday conditions, encompassing 400–800 km in the horizontal and altitudes ranging from 300 to 9,000 m above ground level. Two aircraft flew together, collecting two to four vertical levels (level legs), often stacked one above the other, and typically about 8–12 vertical profiles per flight day. Pal et al. (2020) found that large horizontal and vertical gradients of $[\text{CO}_2]$ exist across frontal boundaries based on the data collected from the summer 2016 campaign. The cross-frontal $[\text{CO}_2]$ contrasts are greatest in the atmospheric boundary layer (ABL), ranging from 5 to 30 ppm, while the contrasts are about 3–5 ppm in the free troposphere (FT). In the vertical dimension, higher $[\text{CO}_2]$ appears in the FT than in the ABL in the cold sector while the opposite pattern appears in the warm sector. Averaged ABL-to-FT $[\text{CO}_2]$ differences can be about 12 and –6 ppm in the warm and cold sectors, respectively. These unique flights were designed to be highly sensitive to the seasonal magnitudes of regional-scale carbon fluxes and to provide broad spatial coverage that cannot be obtained with the current long-term observing network. The third objective is to explore to what degree the aircraft $[\text{CO}_2]$

Table 1
WRF-Chem Model Physics Parameterization Choices

Parameterization	Option used	Reference
Microphysics	Thompson	Thompson et al. (2004)
Longwave radiation	RRTMG longwave scheme	Iacono et al. (2008)
Shortwave radiation	RRTMG shortwave scheme	Iacono et al. (2008)
PBL scheme	MYNN2	Nakanishi and Niino (2006)
Land surface	Unified Noah land-surface model	Chen and Dudhia (2001)
Cumulus	Kain-Fritsch (new Eta) scheme	Kain (2004)

Abbreviations: MYNN2, Mellor-Yamada Nakanishi and Niino Level 2; PBL, planetary boundary layer.

measurements and tower measurements yield consistent evaluations of the modeled fluxes. Merging these two [CO₂] observing systems lends more confidence to our conclusions regarding the modeled fluxes and provides a test of the two observing systems.

The three objectives of this study are addressed through comparing the WRF-Chem [CO₂] simulated using the CASA L2 flux ensemble members to the airborne and tower-based [CO₂] data for the ACT-America summer 2016 campaign period (July 18 to August 28, 2021). The WRF-Chem model setup and methods are described in Section 2. All the data used in this work are described in Section 3. The modeled and observed [CO₂], CASA parameter constraints, and causes of the model errors are illustrated in Section 4. Sections 5 and 6 are discussion and conclusions, respectively.

2. Materials and Methods

2.1. Transport Model Setup

All transport model simulations use WRF-Chem version 3.6.1 (Grell et al., 2005; Skamarock et al., 2008). The modification made to transport greenhouse gases as passive tracers (Lauvaux et al., 2012) allows us to carry ensemble tracers for biogenic CO₂ fluxes (Section 2.2) and [CO₂] boundary conditions (Section 2.4) in one transport run (Section 2.3). In each transport run, WRF-Chem carries 39 CO₂ tracers for the ensembles of boundary conditions and biogenic fluxes, ocean, fossil fuel, and biomass burning fluxes. The modeled total [CO₂] is the sum of a boundary condition, a biogenic flux, and the oceanic flux, fossil fuel emission, and biomass burning CO₂ tracers as described in Feng, Lauvaux, Davies, et al. (2019) and Feng, Lauvaux, Keller, et al. (2019). The [CO₂] boundary condition tracers are propagated into WRF-Chem hourly with the consideration of the conservation of mass (Butler et al., 2020). Five global CO₂ inversion/reanalysis systems are used for [CO₂] boundary conditions (Section 2.4) and 29 biogenic CO₂ fluxes are used for the biogenic CO₂ tracers (Section 2.2). The CO₂ oceanic flux, fossil fuel emission, and biomass burning are taken from CarbonTracker version 2017 (CT2017; Peters et al., 2007).

The same model configurations in Feng, Lauvaux, Davies, et al. (2019) and Feng, Lauvaux, Keller, et al. (2019) are used except for the meteorological initial and boundary conditions. In this study, we used the ERA5 reanalysis (Hersbach et al., 2020) and benchmark the model transport by nudging the WRF-Chem simulation to ERA5. The wind evaluations show that nudging clearly improves model transport (see Text S1 in supporting information). A suite of transport runs is created for uncertainty quantification (Section 2.3). Choices of the model physics schemes are summarized in Table 1. All WRF-Chem simulations have a horizontal resolution of 27 × 27 km for the period from July 18 to August 28, 2016 covering the ACT-America summer 2016 aircraft campaign at hourly resolution. Vertically, WRF-Chem has 50 levels from surface to 50 hPa with 29 levels within 2 km above ground.

2.2. Biogenic CO₂ Flux Ensemble: CASA and CT2017

We include 29 biogenic CO₂ fluxes in each transport run as separate tracers: the 27-member CASA L2 NEE ensemble, the mean of the CASA NEE ensemble, and the CT2017 posterior biogenic CO₂ flux.

Table 2
Perturbations of the Transport Ensemble

Transport run	PBL scheme	Surface layer	LSM	SKEBS
1	MYNN	MYNN	Noah	N
2	MYNN	MYNN	Noah	Y
3	MYNN	MYNN	Noah	Y
4	MYNN	MYNN	Noah	Y
5	MYJ	Eta	RUC	N
6	MYJ	Eta	RUC	Y
7	MYJ	Eta	RUC	Y
8	YSU	MM5	Thermal	N
9	YSU	MM5	Thermal	Y
10	YSU	MM5	Thermal	Y

Abbreviations: LSM, land surface model; MYNN, Mellor-Yamada Nakanishi and Niino; PBL, planetary boundary layer; SKEBS, stochastic kinetic energy backscattering scheme.

The CASA ensemble members (Zhou, Williams, Lauvaux, Feng, et al., 2020) were generated by perturbing the maximum light use efficiency (E_{\max}), optimal temperature of photosynthesis (T_{opt}), and temperature response of respiration (Q_{10}) with the consideration of the biome types in CASA, as described in Zhou, Williams, Lauvaux, Davis, et al. (2020). These three perturbed parameters were determined to dominate the sensitivity of modeled CO_2 fluxes to the model parameters according to an Extended Fourier Amplitude Sensitivity Testing analysis. The initial range for each parameter was broadly sampled for the L1 ensemble. Parameter ranges were subsequently narrowed to those consistent with AmeriFlux data, resulting in a L2 ensemble. CASA simulates gross primary productivity (GPP), total ecosystem respiration (Re), and NEE at monthly resolution. The monthly GPP and Re fluxes were then downscaled to 3-hourly resolution by 3-hourly air temperature and shortwave downward radiation from the North American Regional Reanalysis (Mesinger et al., 2006) using the Olsen and Randerson (2004) method. Two sets of flux products are included in the official release of the CASA flux ensemble, one at 500-m resolution covering the US CONUS region and the other at 5-km covering a broader swath of North America. We used the L2 5-km CASA ensemble for this study. Details about the CASA ensemble products can be found in Zhou, Williams, Lauvaux, Davis, et al. (2020) and Zhou, Williams, Lauvaux, Feng, et al. (2020).

Unlike CASA that directly simulates biogenic CO_2 fluxes (“bottom-up”), CT2017, a “top-down” flux estimate, optimizes the a priori fluxes (CASA; Potter et al., 1993) by assimilating observed $[\text{CO}_2]$ (Peters et al., 2007). CT2017 global 3-hourly posterior biogenic fluxes at $1^\circ \times 1^\circ$ are used in this study. Note that, unlike the CASA ensemble used spun up to equilibrium, the CT2017 biogenic flux prior from CASA include an assumed global terrestrial biospheric sink of 2 PgC/yr.

Both CASA and CT2017 biogenic CO_2 fluxes have a 3-hourly temporal resolution. To downscale to hourly fluxes for the transport model simulations, values in each 3-hourly flux file are repeated hourly over the period they represent.

2.3. Transport Ensemble

The transport ensemble runs are generated using the combination of multiple physical parameterizations and the stochastic kinetic energy backscattering scheme (SKEBS; Berner et al., 2009; Shutts, 2005). Through this combination, model meteorological initial conditions and physics were perturbed at the same time, introducing transport uncertainty due to the model dynamics and physics. Previous studies demonstrated that model errors can be best captured by a combination of multi-physics and SKEBS (Berner et al., 2011, 2015) as opposed to a single perturbation scheme. Feng, Lauvaux, Davies, et al. (2019) and Feng, Lauvaux, Keller, et al. (2019) for the first time applied this combination to simulations of $[\text{CO}_2]$ and demonstrated that a relatively small transport ensemble can represent model $[\text{CO}_2]$ transport uncertainty. We use the root-mean-square deviation (RMSD) of the simulated $[\text{CO}_2]$ transport ensemble from the ensemble mean to represent transport uncertainty. Similar calculations are used to estimate biogenic flux and boundary condition uncertainty.

Here we varied the land surface models (LSMs) and planetary boundary layer (PBL) schemes in WRF-Chem based on the sensitivity study conducted by Díaz-Isaac et al. (2018). The ensemble members are (a) Mellor-Yamada Nakanishi and Niino Level 2.5 (MYNN 2.5) PBL scheme with Noah LSM, (b) Mellor-Yamada-Janjic PBL scheme with RUC LSM, and (c) Yonsei University PBL scheme with five-layer thermal diffusion LSM. The associated surface layer parameterizations are MYNN, Eta, and MM5, respectively. A summary of transport ensemble members can be found in Table 2. Note that these transport ensemble runs are free runs, not nudged to ERA5.

2.4. Boundary Condition Ensemble

A suite of optimized $[\text{CO}_2]$ from five global inversion systems are collected for this study. They are CT2017 (Peters et al., 2007, with updates documented at <http://carbontracker.noaa.gov>), TM5 as described in Basu et al. (2016), GEOS-Chem developed in the Carbon Monitoring System (Liu et al., 2014), GEOS-Chem as described in Schuh et al. (2019), and PCTM as described in Barker et al. (2004). These global modeled $[\text{CO}_2]$ fields are propagated into WRF-Chem hourly as separate tracers following Butler et al. (2020).

2.5. Footprint Analysis

To understand the fluxes influencing the observations used in this study, the Lagrangian Particle Dispersion Model (Uliasz et al., 1994) is used to create surface influence functions (footprints) for each flight in the Summer 2016 ACT aircraft campaign (see description in Section 3.1). Particles are released along the time of each aircraft transect within the boundary layer, and are traced backwards in time over a two-week period using meteorology provided by the WRF nudged-transport simulation. The influence by the surface is represented by the number of particles interacting with the surface grid, that is, below 50 m above ground, are summed up, providing a temporal and spatial function that relates the signal observed by the aircraft to the surface fluxes responsible for that signal (Seibert & Frank, 2004).

In addition to aircraft $[\text{CO}_2]$ measurements, we also use a subset of the ground-based $[\text{CO}_2]$ tower measurements from the NOAA ObsPack GlobalViewPlus package (Cooperative Global Atmospheric Data Integration Project, 2019; see Section 3.2). We create footprints for the towers by releasing particles at 21 UTC backward in time over two-week period using the nudged transport from July 18 to August 28, 2016 across 23 different tower sites in the US. The selection of 21 UTC limits the tower observations to well-mixed ABL conditions, minimizing model transport uncertainty.

2.6. Model Evaluation Metrics

To demonstrate the importance of the uncertainty in the biosphere, we first compare the contribution of biosphere, transport, and boundary conditions to the modeled $[\text{CO}_2]$ uncertainty. We use the RMSD of a given component to illustrate the ensemble spread and associated uncertainty.

We focus on the performance of the individual biogenic CO_2 flux members through the biome-based model biases, which is determined by the CASA biome map and the footprint of the measurements. The uncertainties in modeled $[\text{CO}_2]$ associated with the individual biogenic flux members are determined by the spread of the transport and boundary conditions ensembles.

To investigate the causes of modeled $[\text{CO}_2]$ biases and eliminate the potential interference of transport and boundary conditions inherent in $[\text{CO}_2]$ comparisons, we also directly compare the CASA flux members and CT2017 fluxes member with AmeriFlux eddy-covariance flux measurements. This comparison tests for consistency between $[\text{CO}_2]$ and eddy-covariance flux analyses and lends more insight into the diel cycle of CO_2 fluxes. We group these analyses according to the dominant biomes surrounding each $[\text{CO}_2]$ tower.

Due to the uneven distribution of flux and $[\text{CO}_2]$ towers, flux towers cannot be always found in the footprint of a given $[\text{CO}_2]$ tower. We therefore propagate the flux biases at each $[\text{CO}_2]$ tower location using Equation 1 below. By assuming that a given biome type in the CASA model and CT2017 have a similar behavior (bias) everywhere, we group the flux towers influenced by the same dominant biome together and calculate the overall flux, F_T , and bias for the given biome. The flux biases at a given $[\text{CO}_2]$ tower location B_i , can be expressed in the following equation.

$$B_i = \frac{1}{24} \sum_{t=1}^{24} \left(\sum_{T=0}^{14} F_T W_T \right) W_t, \quad (1)$$

where W_T and W_t are the spatial and temporal weighting functions that propagate the flux biases from the biome level to the $[\text{CO}_2]$ site level, respectively. The subscript T denotes the biome: Indexes 1 to 14 are the biome in CASA following Zhou, Williams, Lauvaux, Davis, et al. (2020); index 0 represents water bodies. W_T are the fractional areas of individual biomes relative to the entire summer-averaged influence area of

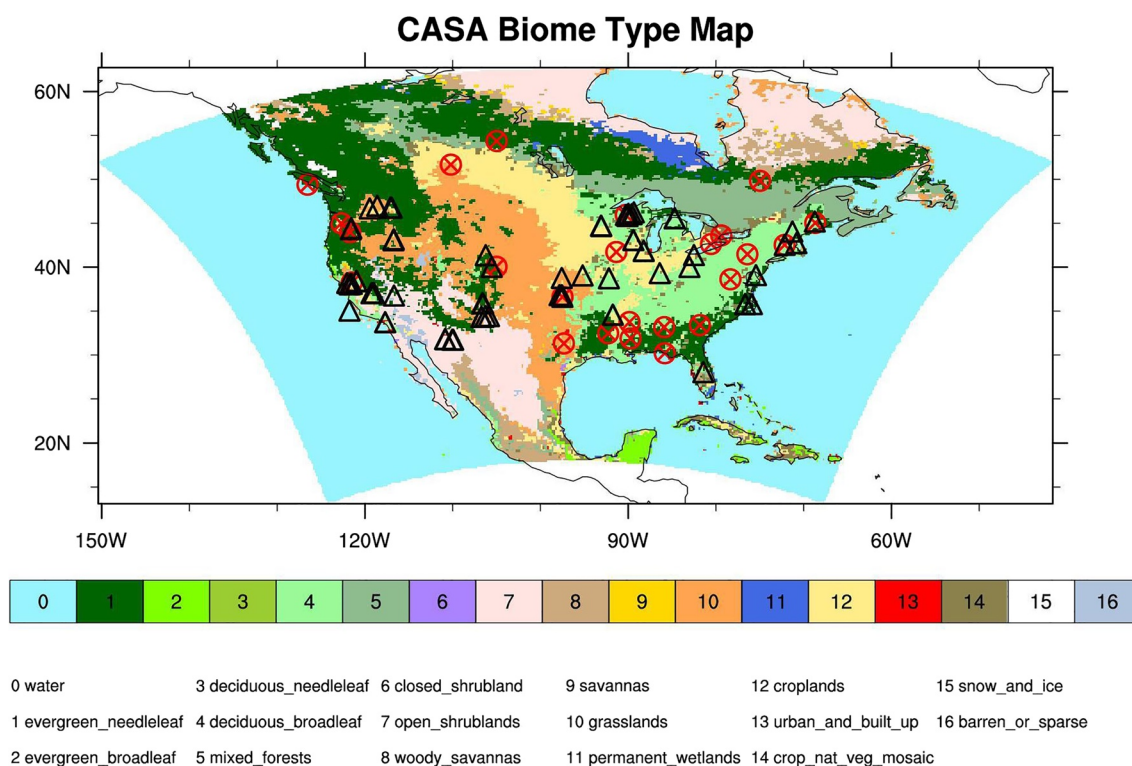


Figure 1. Carnegie-Ames-Stanford-Approach (CASA) biome type regridded to the 27×27 -km WRF-Chem grids, and the locations of the $[\text{CO}_2]$ towers (red circled crosses) and AmeriFlux flux towers (black triangles). The colored area is the model domain.

tower-based, daytime $[\text{CO}_2]$ observations (Figure S2). In the analysis, we find that some towers can be influenced by up to six biome types according to the footprint of the tower $[\text{CO}_2]$ measurements. The temporal weighting function, W_t , is essentially equal across the 24 h of a day due to the indistinguishable diel cycle after averaging the footprint over the period of interest (Figure S3b and Text S2 in SI). We acknowledge that a more thorough analysis should consider day-to-day variation in W_t for each tower observation. We hypothesize that the day-to-day variation in W_t will have minimal impact on the bias analysis we focus on in this study.

3. Data

3.1. ACT-America Aircraft Data

Two aircraft, the NASA Langley Beech-craft B200 King Air and NASA Goddard Space Flight Center's C-130H Hercules aircraft, were used to collect high quality in situ measurements of greenhouse gases, other gas species, and meteorological fields over Mid-Atlantic (MA), Mid-West (MW), and South Gulf (South) regions of the United States. The flight dates and patterns can be found Table S1. All the flights during this time took off around noon and landed around 5 p.m. local time. In this study, we used ACT-America L3 Merged in situ Atmospheric Trace Gases and Flask Data (Davis et al., 2018). This product provides integrated measurements and metadata flag information, including flight pattern, air mass type, and boundary layer information at five-second intervals. More information about the ACT-America campaign and measurements can be found at <https://actamerica.ornl.gov/> and Davis et al. (2021). The nearest point interpolation is applied to extract modeled $[\text{CO}_2]$ along the flight tracks.

3.2. NOAA ObsPack GlobalViewPlus $[\text{CO}_2]$ Product

We also use tower-based in situ $[\text{CO}_2]$ data from the NOAA ObsPack GlobalViewPlus product (Cooperative Global Atmospheric Data Integration Project, 2019; see locations in Figure 1). Twenty-three tower locations

Table 3
Selected [CO₂] Towers and Top 2 Biome Types

Site	Lat	Lon	Intake elevation (m ASL)	1st biome type (fraction %)	2nd biome type (fraction %)
AMT	45.03	-68.68	160	DB (42.1)	EN (32.2)
BAO	40.05	-105	1,884	GL (62.8)	EN (36.4)
CPS	49.82	-74.98	389	EN (46.3)	MF (32.2)
ESP	49.38	-126.54	47	EN (49.6)	Wa (47.1)
ETL	54.35	-104.99	597	EN (67.8)	CR (14.0)
GCI01	32.47	-92.28	165	EN (61.2)	DB (22.3)
GCI02	33.75	-89.85	205	DB (56.2)	CR (27.3)
GCI03	31.89	-89.73	232	EN (57.9)	DB (33.9)
GCI04	33.18	-85.89	428	EN (58.7)	DB (40.5)
GCI05	30.2	-85.83	105	Wa (41.3)	EN (41.3)
HFM	42.54	-72.17	369	DB (87.6)	Wa (10.0)
HNP	43.61	-79.39	97	CM (43.8)	DB (17.4)
LEF	45.95	-90.27	868	DB (81.8)	Wa (7.4)
MBO	43.98	-121.69	2742	EN (73.6)	GL (20.6)
MRC	41.47	-76.42	652	DB (98.3)	CM (1.7)
OSI	45	-122.7	620	EN (74.4)	Wa (12.4)
SCT	33.41	-81.83	420	EN (78.5)	DB (17.4)
SGP	36.61	-97.49	374	GL (77.7)	DB (12.4)
SNP	38.62	-78.35	1,025	DB (93.4)	EN (3.3)
TPD	42.64	-80.56	266	DB (32.2)	CR (27.3)
WBI	41.72	-91.35	621	CR (76.9)	DB (17.4)
WGC	38.27	-121.49	483	DB (26.4)	CR (24.0)
WKT	31.31	-97.33	708	GL (67.8)	DB (22.3)

Abbreviations: CM, cropland natural vegetation mosaic; CR, croplands; DB, deciduous broadleaf forest; EN, evergreen needleleaf forest; GL, grasslands; MF, mixed forest; Wa, water.

(Table 3) from this data package were selected for investigations. The ObsPack product collects greenhouse gas data from providers around the globe and reformats the data into the ObsPack framework in support of carbon cycle modeling studies (Masarie et al., 2014). The [CO₂] data are first organized into hourly data, and then nearest point interpolation is applied to extract modeled [CO₂] at the tower locations.

3.3. AmeriFlux CO₂ Flux Measurements

We also include CO₂ flux measurements to evaluate our findings regarding biogenic CO₂ flux members. The CO₂ flux data are obtained from the eddy-covariance measurements from the AmeriFlux network (<https://ameriflux.lbl.gov>). Seventy-one flux tower sites from the domain of interest were used. The data providers are responsible for data quality control. The locations and information of the sites can be found in Figure 1 and Table S2. We obtained a single estimate of NEE for each flux tower location from the non-gap-filled NEE values reported by AmeriFlux, with a preference for eddy-covariance measurements with a storage correction. Tower-measured NEE was averaged to three-hour intervals to match the time resolution of CASA and CT2017. No intervals were excluded if they had any reported data.

4. Results

4.1. Spatiotemporal Variability of [CO₂]

We select five of the 25 research flights (Figure 2) to illustrate the typical flight patterns for fair and frontal weather regimes. As expected, both models and observations show large [CO₂] gradients in the ABL, leading to large variations in [CO₂bio]. Pal et al. (2020) reported that an elevated [CO₂] band was repeatedly observed along the cold frontal boundary, a feature also captured by the simulations. The 8/4/2016 frontal case shows a narrow, elevated [CO₂] band at the frontal boundary (~25 ppm difference in [CO₂] across the front). We examined this frontal case with a 3 × 3-km (cloud-resolving) resolution model and showed that this elevated [CO₂] band has a maximum width of ~200 km and a length of over 800 km extending from northeastern Kansas to northeastern Iowa (Samaddar et al., 2021). Figure 2 shows that the frontal boundaries are

associated not only with elevated [CO₂] but also with highly variable [CO₂bio]. More than 5 ppm RMSD of [CO₂bio], caused by variability among the CASA ensemble fluxes, appears in the ABL along the frontal boundaries. Enhanced by the baroclinic instability, the cold airmass on the west lifted the warm airmass aloft. ABL [CO₂] penetrates into the free troposphere along the frontal lifting. RMSD in [CO₂bio] greater than 3 ppm, an indicator of strong surface influence, reaches up to 3.5 km above sea level on 7/18 and 1.7 km on 8/4 (Figure 2). On the contrary, both [CO₂] and [CO₂bio] have less variability for the fair-weather cases.

We use the ABL [CO₂] observations to evaluate biogenic CO₂ fluxes since, as Figure 2 illustrates, these data are the most sensitive to variations in the biological CO₂ fluxes. We limit our work to ABL [CO₂] for the rest of this analysis. Figure 3 shows the averaged [CO₂] sampled by aircraft and simulations. Note that the modeled [CO₂] are from the transport-nudged simulation (described in Section 2.1), while the uncertainties are determined by spreads of the ensemble runs associated with different components described in Sections 2.2, 2.3 and 2.4. All ACT-America aircraft collected afternoon samples aiming at well-mixed ABL conditions. The most outstanding feature overall is that the biosphere is the most uncertain component in the simulation for the fair and frontal weather regimes except for three cases in the South and one in the Midwest. The footprint analysis (Figure S4) shows that the Southern flights are mainly influenced by

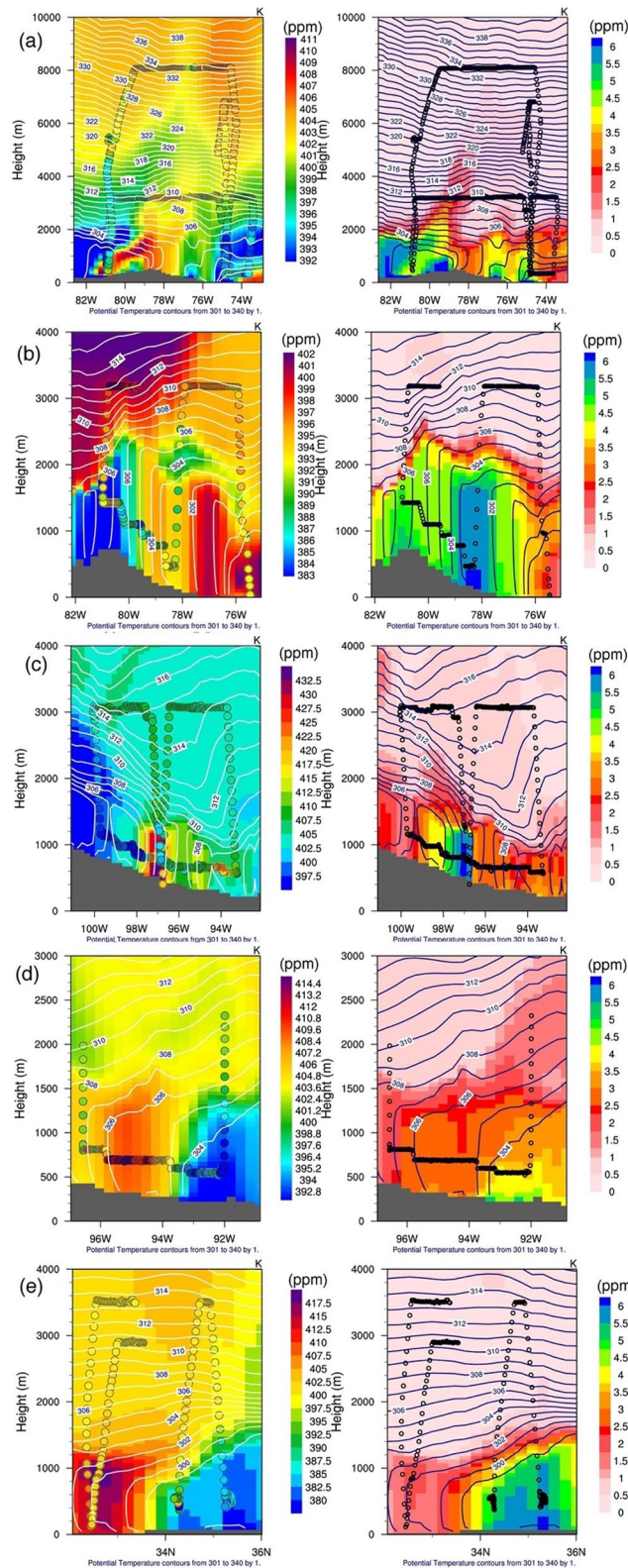


Figure 2. Simulated $[CO_2]$ (left column) and the root-mean-square deviation (RMSD) of the biogenic $[CO_2]$ ($[CO_2]_{bio}$) from the mean of modeled $[CO_2]_{bio}$ (right column) for five flights in the ACT Summer 2016 aircraft campaign: (a) 7/18 19 UTC (front in MA); (b) 7/21 19 UTC (fair Wx in MA); (c) 8/4 18 UTC (front in MW); (d) 8/9 19 UTC (fair Wx in MW); (e) 8/22 22 UTC (fair Wx in South). The aircraft sampled $[CO_2]$ are overlaid with the simulations (left column). The black dots on the RMSD panels are the flight paths. Potential temperature contours are overlaid to indicate the frontal locations.

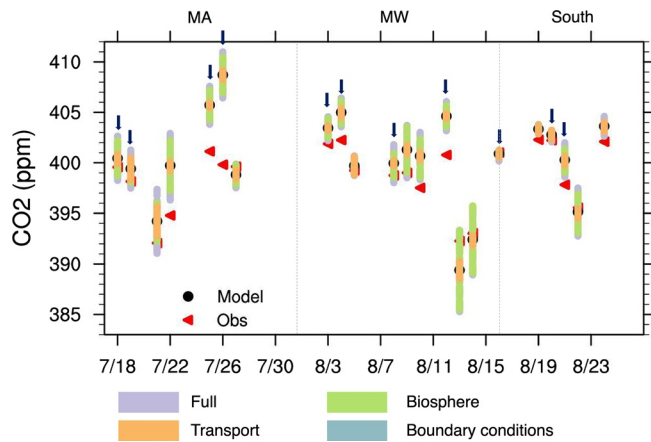


Figure 3. Daily afternoon averaged $[CO_2]$ from the ACT aircraft and simulations. The error bars of the simulated $[CO_2]$ are the spread of the full, biogenic flux, transport, and boundary condition ensembles. Note that the spread of boundary condition is very small and not visible in the figure. Only the atmospheric boundary layer samples are included in the analysis. The point-down arrows denote the frontal cases during the campaign. Gray dashed lines denote the transit flights. “MA,” “MW,” and “South” on the top of the figure denote the Mid-Atlantic, Mid-West, and South of Gulf Coast flights.

Gulf onshore flow and, thus had a limited flux footprint over land. The one Midwestern flight was an OCO-2 underflight in the western Dakotas where biological fluxes are small.

Observations tend to be encompassed by the ensemble model spread except on four flight days: 7/22, 7/25, 7/26, and 8/12, on which the model shows large discrepancy with observations. Our preliminary investigation indicates that the disagreements on 7/25 and 7/26 are mainly caused by unrealistically strong uptake to the west of the Appalachia area in the CT2017 biogenic fluxes; 8/12 is due to errors in the long-range transport.

In the following sections, we focus on investigating the coherence between aircraft and tower $[CO_2]$ data and evaluating the model biases across the CASA ensemble members. CT2017 serves as the reference for this exercise. As an inversion product that is constrained by atmospheric data, we expect that CT2017 should agree well with observed $[CO_2]$ even though the ACT-America aircraft data were not assimilated in CT2017. Figure 4 shows the model skill in the $[CO_2]$ simulations driven with different biogenic CO_2 fluxes. Note that both aircraft and tower $[CO_2]$ data are kept at native temporal resolutions in the comparisons, reflecting the model performance in capturing the spatiotemporal variability of $[CO_2]$ at the synoptic scale (no diel variation due to only afternoon samples used). In general, both aircraft and tower $[CO_2]$ comparisons show that skill of the CASA members is similar to CT2017 in capturing the variability in ABL $[CO_2]$ as indicated by the similar correlations with observations, but better in capturing the magnitude of this variability (normalized stand-

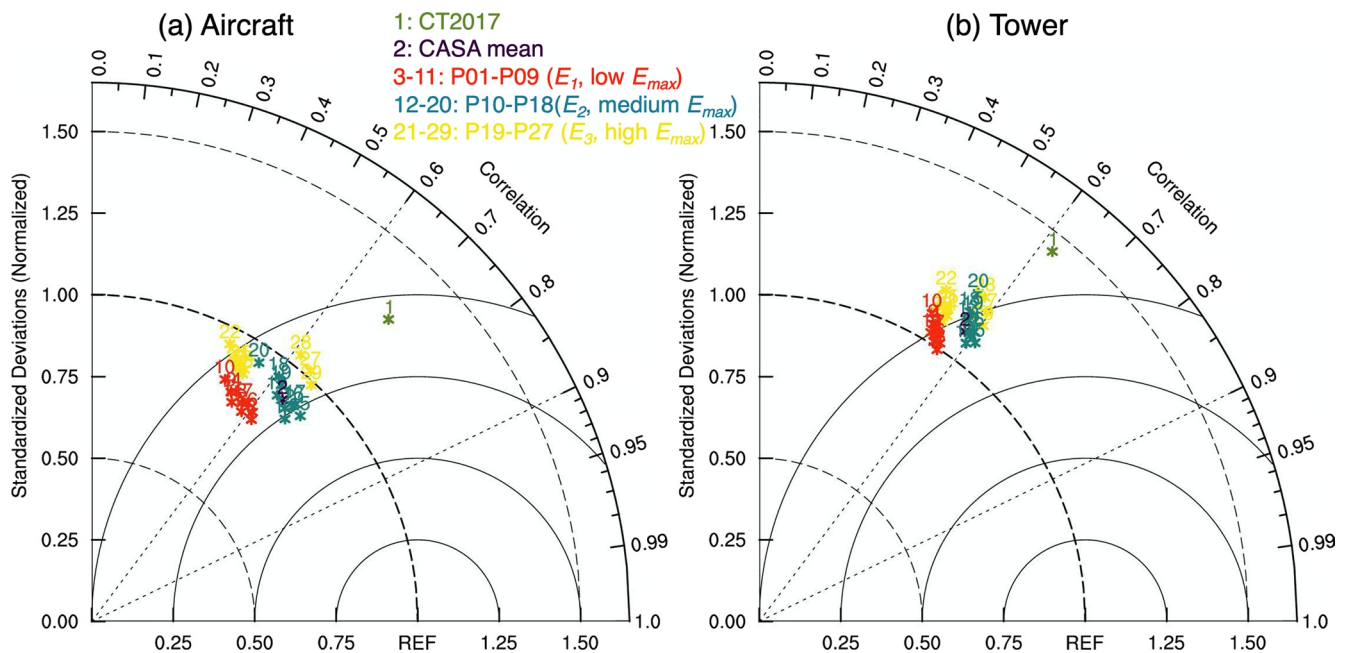


Figure 4. Taylor diagrams of simulated $[CO_2]$ compared with the observations. The standard deviations of models are normalized by the standard deviations of observations (“REF”). The model $[CO_2]$ are associated with the (29) biogenic flux members from Carnegie-Ames-Stanford-Approach (CASA) and CarbonTracker version 2017 (CT2017). The color scheme indicates CT2017, CASA ensemble mean, CASA low E_{max} (E_1), CASA medium E_{max} (E_2), and CASA high E_{max} (E_3) groups. “P#” in the legend denotes the index of CASA ensemble member in the original product. The observations used in (a) are from the atmospheric boundary layer legs of the ACT-America aircraft measurements. The flight dates and other information are listed in Table S1. The observations used in (b) are from the $[CO_2]$ tower measurements at afternoon hours (19–22 UTC). The locations and information can be found in Figure 1 and Table 3. Both aircraft and tower $[CO_2]$ data are kept at native temporal resolutions. Modeled $[CO_2]$ are determined by the value of the nearest grid cell to sample locations. The period of interest is from July 18 to August 28, 2016 covering the ACT-America summer 2016 aircraft campaign.

Table 4
Top 2 Biome Types in the ACT-America Summer 2016 Flight Samples

Date	First biome type (fraction %)	Second biome type (fraction %)
2016-07-18 ^a	DB (60.1)	CR (17.4)
2016-07-19	DB (40.7)	EN (20.7)
2016-07-21 ^a	DB (54.6)	Wa (15.7)
2016-07-22	DB (60.7)	EN (16.8)
2016-07-25	DB (69.0)	EN (16.4)
2016-07-26	DB (61.0)	CR (17.6)
2016-07-27	DB (30.1)	EN (19.1)
2016-08-03	GL (38.0)	OS (24.5)
2016-08-04 ^a	GL (40.3)	Wa (25.5)
2016-08-05	GL (40.9)	Wa (17.5)
2016-08-08	DB (34.6)	CR (24.1)
2016-08-09 ^a	DB (44.0)	GL (21.6)
2016-08-10	DB (35.7)	CR (20.1)
2016-08-12	Wa (25.1)	CR (20.1)
2016-08-13	CR (72.6)	GL (12.0)
2016-08-14	CR (62.7)	EN (9.0)
2016-08-16	Wa (71.3)	DB (9.4)
2016-08-19	Wa (93.3)	DB (3.3)
2016-08-20	Wa (66.2)	CR (14.1)
2016-08-21	CR (42.4)	Wa (22.0)
2016-08-22 ^a	CR (41.3)	DB (28.4)
2016-08-24	Wa (83.9)	EN (5.2)

Abbreviations: ACT, Atmospheric Carbon and Transport; CR, croplands; DB, deciduous broadleaf forest; EN, evergreen needleleaf forest; GL, grasslands; OS, open shrublands; Wa, water.

^aDenotes the flights shown in Figure 2.

ardized deviations). CT2017 overestimates the magnitude of variability in ABL [CO₂]. Tower and aircraft yield similar results.

Aligning the aircraft and tower comparisons together, the CASA-simulated [CO₂] tend to somewhat underestimate the variability of aircraft [CO₂] (root-mean-squared errors) but overestimate that of tower [CO₂]. Given the different sampling strategies between aircraft and tower measurements, we hypothesize that the models have a tendency to underestimate the spatial variability of atmospheric [CO₂] but overestimate temporal variability at the synoptic scale.

Another feature worth noting is that the CASA-simulated [CO₂] are clustered into two groups. Some high E_{\max} members fall into the low E_{\max} group, and the rest falls in the medium E_{\max} group. This grouping is evident in both aircraft and tower [CO₂] comparisons. Given the same transport and boundary conditions were used, the clustering is likely driven by the difference in biogenic fluxes associated with the values of the three parameters used for generating the CASA flux ensemble.

4.2. Identification of the CASA Key Parameters

The modeled [CO₂] biases associated with individual CASA flux members and CT2017 with the uncertainty bounds that are determined by the transport and boundary condition ensembles are shown in Figures S6 and S7 for each observation from aircraft and tower comparison, respectively. Applying the footprint analysis to aircraft and tower measurements and the CASA-defined biome map, we obtain the model biases for each biome. We list the top 2 biomes with the most influence on each aircraft and tower measurements in Tables 3 and 4, respectively. Croplands (CR), deciduous broadleaf forest (DB), and grasslands (GL) are the major biome types sampled by aircraft; CR, DB, GL, and evergreen needleleaf forest (EN) are mainly sampled by towers (Figure 5). The samples primarily influenced by water bodies are removed from the results. Two aspects stand out in the comparisons, modeled [CO₂] from all the flux members are positively biased across all biome types, and one group of the CASA flux members has better agreement with the observations than others. These two aspects are illustrated consistently in the comparisons of both aircraft and tower measurements.

For the first aspect, we find that all flux estimates, including CT2017 and CASA ensemble, overestimate [CO₂]. This is also reflected in Figure 3. Figure 3 also displays that both biogenic flux and transport can play an important role in modeled [CO₂] errors. We first focus on the relative performance across the ensemble members to isolate the influence from model transport. The model biases seem scaled with the degree of plant productivity, given that we find larger biases associated with CR and DB and smaller biases associated with EN and GL. Zhou, Williams, Lauvaux, Davis, et al. (2020) also reported that the monthly averaged NEE of the CASA ensemble averaged over 13 years had a larger positive bias in CR and DB than EN and GL in summer. When comparing aircraft-based biases to tower-based biases, the WRF-Chem simulated [CO₂] is more positively biased. Assuming that spatial variability in [CO₂] is related to temporal variability, as towers and aircraft basically observe the same weather systems, the larger aircraft-based biases might be caused by different geographic sampling or by the enhanced variability in aircraft data (designed to sample frontal systems) leading to larger biases. As expected, CT2017 has better agreement with the observations overall since its fluxes have been optimized using atmospheric [CO₂] data. However, the fact that WRF-Chem performs better when coupled to optimized CT2017 biogenic CO₂ fluxes confirms that transport model differences remain much smaller than flux differences.

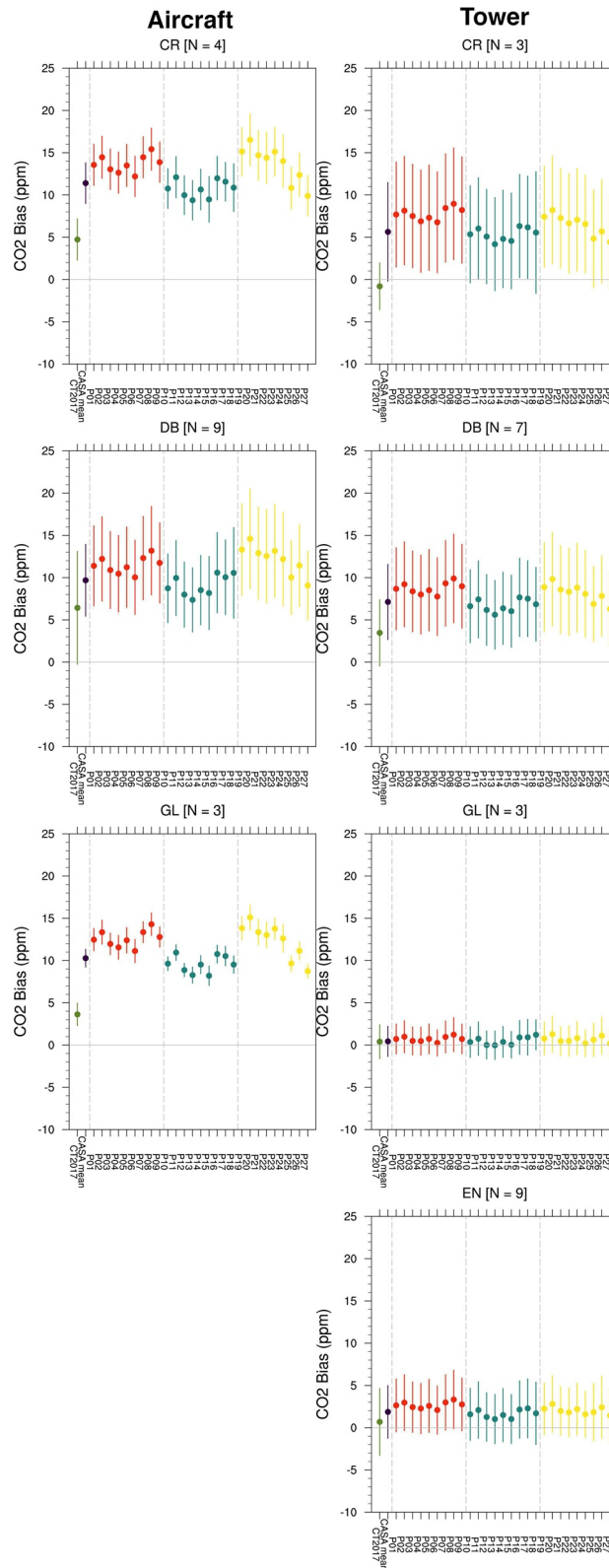


Figure 5. Modeled $[CO_2]$ biases against aircraft and tower measurement during afternoon hours (from noon-6 p.m. local time) on the basis of different biomes for the period from July 18 to August 28, 2016. The model $[CO_2]$ are associated with the (29) biogenic flux members from Carnegie-Ames-Stanford-Approach (CASA) and CarbonTracker version 2017 (CT2017). The color scheme indicates CT2017, CASA ensemble mean, CASA low E_{max} (E_1), CASA medium E_{max} (E_2), and CASA high E_{max} (E_3) groups. “P#” on the x-axis are the indexes of CASA ensemble members in the original product.

Ranking	CR		DB*		GL		EN*	E1: Low E _{max} E2: Medium E _{max} E3: High E _{max}
	Aircraft (N=4)	Tower (M=3)	(N=9)	Tower (M=7)	Aircraft (N=3)	Tower (M=3)	Tower (M=9)	
1	CT2017	CT2017	CT2017	CT2017	CT2017	P12/224	CT2017	
2	P13/022	P13/022	P13/022	P13/022	P15/222	P13/022	P13/022	
3	P15/222	P27/236	P12/224	P15/222	P13/022	P15/222	P15/222	
4	P27/236	P15/222	P15/222	P12/224	P27/236	P27/236	P12/224	
5	P12/224	P14/-222	P14/-222	P27/236	P12/224	P24/232	P27/236	
6	P14/-222	P25/036	P10/024	P14/-222	P18/226	P06/212	P14/-222	
7	P10/024	P12/224	P27/236	P10/024	P14/-222	P10/024	P10/024	
8	P25/036	P10/024	CASA mean	P18/226	P10/024	P14/-222	P24/232	*E2 = E3 for DB and EN
9	P18/226	P18/226	P11/-224	P25/036	P25/036	CT2017	P18/226	
10	CASA mean	CASA mean	P25/036	CASA mean	CASA mean	CASA mean	P22/032	
11	P17/-226	P26/-236	P06/212	P11/-224	P17/-226	P21/234	P25/036	
12	P16/026	P11/-224	P17/-226	P17/-226	P16/026	P04/012	CASA mean	
13	P11/-224	P17/-226	P04/012	P16/026	P11/-224	P22/032	P21/234	
14	P06/212	P16/026	P18/226	P06/212	P06/212	P03/214	P11/-224	
15	P26/-236	P24/232	P16/026	P26/-236	P26/-236	P25/036	P06/212	
16	P04/012	P22/032	P03/214	P04/012	P04/012	P01/014	P16/026	
17	P03/214	P06/212	P05/-212	P24/232	P03/214	P09/216	P23/-232	
18	P05/-212	P04/012	P01/014	P22/032	P05/-212	P05/-212	P19/034	
19	P01/014	P23/-232	P26/-236	P03/214	P01/014	P11/-224	P04/012	
20	P09/216	P21/234	P09/216	P05/-212	P24/232	P19/034	P17/-226	
21	P24/232	P05/-212	P24/232	P21/234	P09/216	P23/-232	P26/-236	
22	P22/032	P19/034	P02/-214	P01/014	P22/032	P16/026	P03/214	
23	P02/-214	P03/214	P07/016	P23/-232	P02/-214	P17/-226	P05/-212	
24	P07/016	P01/014	P22/032	P19/034	P21/234	P07/016	P01/014	
25	P21/234	P02/-214	P21/234	P09/216	P07/016	P02/-214	P09/216	
26	P23/-232	P20/-234	P23/-232	P02/-214	P23/-232	P26/-236	P20/-234	
27	P19/034	P09/216	P08/-216	P07/016	P19/034	P18/226	P02/-214	
28	P08/-216	P07/016	P19/034	P20/-234	P08/-216	P08/-216	P07/016	
29	P20/-234	P08/-216	P20/-234	P08/-216	P20/-234	P20/-234	P08/-216	

Figure 6. The final ranking of the modeled [CO₂] bias among the biogenic flux members shown in Figure 5. The model skills are ranked decreasingly, meaning the model biases increase from top to bottom. The name convention of the Carnegie-Ames-Stanford-Approach (CASA) ensemble members follows P#/T_{opt}E_{max}Q₁₀. # is the index of the CASA members in the original product. At the position of T_{opt}, “0” means the default value; “2” means the default value plus 2°C; “-2” means the default value minus 2°C. For E_{max}, “1,” “2,” and “3” denotes the low, medium, and high E_{max} in Table 4 of Zhou, Williams, Lauvaux, Davis, et al. (2020) and Zhou, Williams, Lauvaux, Feng, et al. (2020). “2,” “4,” “6” at the Q₁₀ position indicate that Q₁₀ = “1.2,” “1.4,” and “1.6,” respectively.

For the second aspect, we discover that distinct groups of CASA members reflect the three parameter values for maximum light use efficiency, E_{max}. The flux members that show the best agreement with the observations mostly have medium E_{max} values while the groups with low and high E_{max} values correspond to larger model biases. Both aircraft and tower measurements identify that E_{max} is the dominant parameter in the CASA ensemble, which is consistent with the sensitivity results of Zhou, Williams, Lauvaux, Davis, et al. (2020).

The performance of the individual flux members is summarized by ranking them as a function of bias (Figure 6). Medium E_{max} (E₂) leads to better modeled [CO₂] across different biome types and observation platforms. The groups with low (E₁) and high (E₃) E_{max} tend to have similar biases. A few members with high E_{max} are ranked high (in the top one third), such as P27 and P25 for CR, P27 for DB, P27 for GL, and P27 for EN. For DB and EN, both E₂ and E₃ are assigned the same value (medium E_{max}) in CASA perturbation, explaining why some model members with high E_{max} value are ranked high. However, due to the impact of the long-range transport over upwind biomes (where E₂ not equal E₃), the model [CO₂] bias ranking is different for the DB’s and EN’s E_{max} groups.

Further investigating the top-performing 11 ensemble members according to Q₁₀ values (Figure 7), we find that in general, the flux members with low Q₁₀ value (Q₁₀ = 1.2) are ranked high, followed with medium Q₁₀ (Q₁₀ = 1.4), and then high Q₁₀ value (Q₁₀ = 1.6). In contrast, no T_{opt}-driven grouping is visible (not shown). We conclude that Q₁₀ plays a secondary role in CASA-simulated summer NEE.

Ranking	CR		DB*		GL		EN*
	Aircraft (N=4)	Tower (M=3)	Aircraft (N=9)	Tower (M=7)	Aircraft (N=3)	Tower (M=3)	Tower (M=9)
1	CT2017	CT2017	CT2017	CT2017	CT2017	P12/224	CT2017
2	P13/022	P13/022	P13/022	P13/022	P15/222	P13/022	P13/022
3	P15/222	P27/236	P12/224	P15/222	P13/022	P15/222	P15/222
4	P27/236	P15/222	P15/222	P12/224	P27/236	P27/236	P12/224
5	P12/224	P14/-222	P14/-222	P27/236	P12/224	P24/232	P27/236
6	P14/-222	P25/036	P10/024	P14/-222	P18/226	P06/212	P14/-222
7	P10/024	P12/224	P27/236	P10/024	P14/-222	P10/024	P10/024
8	P25/036	P10/024	CASA mean	P18/226	P10/024	P14/-222	P24/232
9	P18/226	P18/226	P11/-224	P25/036	P25/036	CT2017	P18/226
10	CASA mean	CASA mean	P25/036	CASA mean	CASA mean	CASA mean	P22/032
11	P17/-226	P26/-236	P06/212	P11/-224	P17/-226	P21/234	P25/036

*E2 = E3 for DB and EN

Emax = E2* && Q10 = 1.2	Emax = E2* && Q10 = 1.4	Emax = E2* && Q10 = 1.6	Others
-------------------------	-------------------------	-------------------------	--------

Figure 7. Top 11 biogenic flux members in Figure 6 applied different color schemes in order to reveal the secondary parameter, Q_{10} . Note that the flux members associated in the medium E_{max} are with color background; others are with white background.

In summary, the aircraft and tower data deliver consistent results. Given the multiple level-leg sampling and profiling strategies available in the aircraft data in addition to the large vertical gradients in $[CO_2]_{bio}$ that appear in data set (Figure 2), there is potential to impose additional constraints on biogenic fluxes and mixing heights using the vertical gradients of $[CO_2]$. Although outside the scope of this study, these vertical gradients will be examined in future work.

4.3. The Causes of Modeled $[CO_2]$ Biases

The model-tower comparisons show the modeled $[CO_2]$ biases for each $[CO_2]$ tower we included in the previous analysis in Figure 8a. All the CASA members lead to an overestimate of $[CO_2]$ except at three towers: BAO, ETL, and OSI. The CASA flux members result in modeled $[CO_2]$ that is similar to CT2017 for GL and EN sites with the exception of three sites in the South Gulf region (GCI01, GCI03, and GCI04). Variability in the TBM parameters have more impact on CR and DB where they have the larger sink strength and plant productivity. A similar pattern of biases across biome types can be seen in CT2017, though CT2017 shows slightly better agreement with the observations.

Model transport, boundary condition, and other CO_2 flux components also contribute to modeled $[CO_2]$ errors. Although we illustrated that the contribution is less than that from the biosphere, it can potentially shift the biases uniformly up or down. Therefore, to root out the potential interference, following Equation 1, we calculate flux biases associated with each $[CO_2]$ tower over the same time period following Equation 1 (Figure 8b). Consistent with the $[CO_2]$ analysis, all the flux members show positive biases across the $[CO_2]$ towers, indicating the positive biases in modeled $[CO_2]$ are mainly due to the fact that CASA and CT2017 underestimate net uptake. For CASA, this can be attributed to weak plant productivity and/or strong respiration. Zhou, Williams, Lauvaux, Davis, et al. (2020) pointed out that there is a missing net sink due to the lack of crop harvest and forest recovery in the CASA model, yielding a net overestimate in annual respiration. Additionally, the relative performance among the CASA E_{max} groups and CT2017 with respect to flux tower data is consistent with the comparison to $[CO_2]$. The low E_{max} values lead to the largest biases in both $[CO_2]$ and flux space, and CT2017 shows smaller biases for CR and DB. Such coherence across measurement space (mole fraction and flux) and platforms (aircraft, $[CO_2]$ towers, and flux towers) lends a high degree of confidence to these results.

We break the daytime and nighttime flux biases apart to explore what causes the overall weak NEE uptake in CASA. In summer, the daytime fluxes are a combination of GPP and ecosystem respiration; the nighttime fluxes are driven completely by respiration. In the flux observations, the all time, daytime, and nighttime averaged fluxes range from -2 to $-1 \mu mol m^{-2} s^{-1}$, from -8 to $-3 \mu mol m^{-2} s^{-1}$, and from 0.5 to $2.5 \mu mol m^{-2} s^{-1}$, respectively (Figure S5 in SI). Temporal downscaling of CASA's monthly fluxes resulted in consistent under-

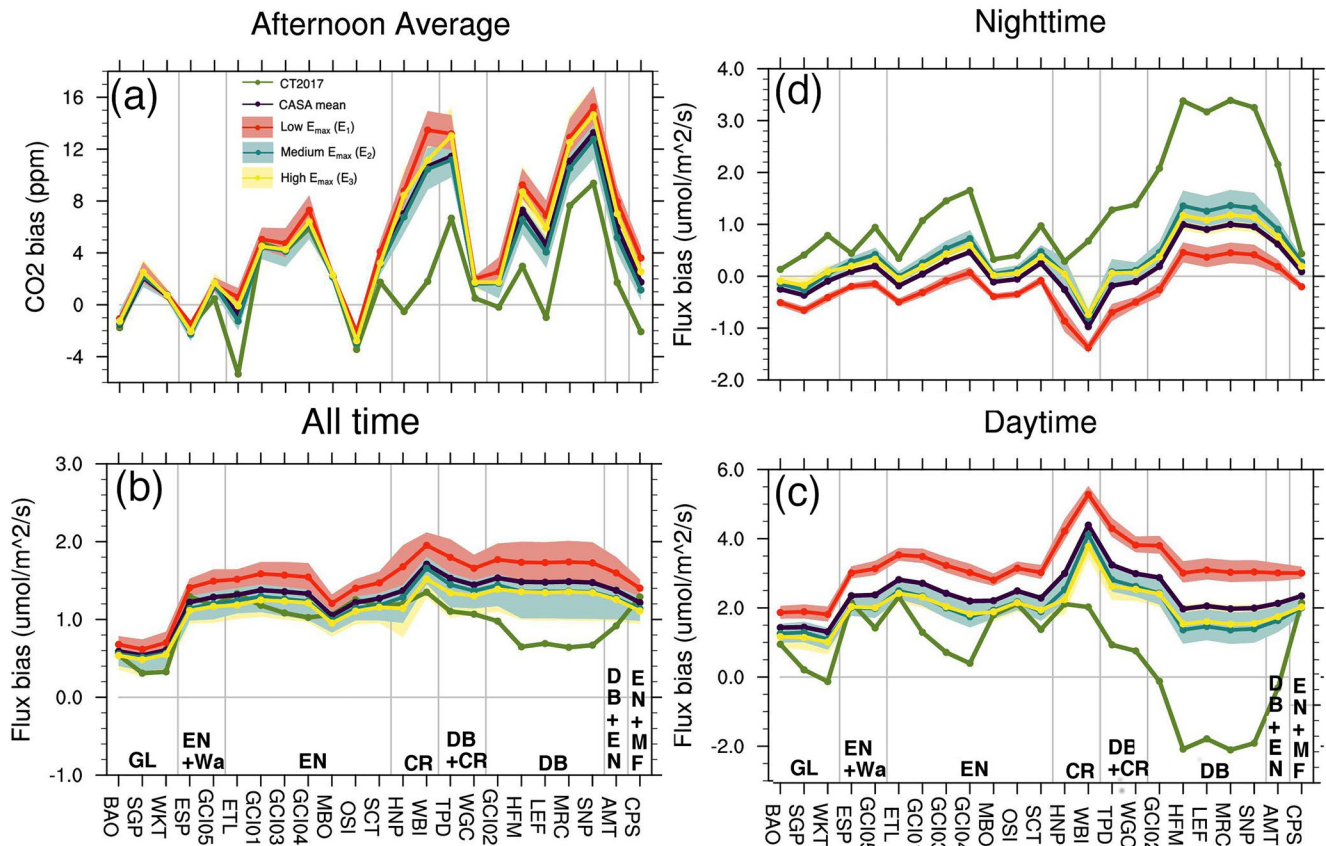


Figure 8. CO₂ mole fraction and flux biases at the [CO₂] towers from July 18 to August 28, 2016. (a) Afternoon averaged (19–22 UTC) [CO₂] biases. (b) All time, (c) daytime (15–2 UTC), and (d) nighttime (3–14 UTC) averages of the flux biases derived from Equation 1. The shaded areas are the minimum and maximum biases among different E_{\max} groups in the Carnegie-Ames-Stanford-Approach (CASA) ensemble. The associated solid lines are the mean of the given E_{\max} group. Additionally, CarbonTracker version 2017 and the CASA ensemble mean are denoted in green and purple dotted lines. The horizontal gray line is unbiased. The vertical gray lines divide the towers into various biomes determined by the footprint analysis. The dominate biomes, GL—grassland, EN—evergreen needleleaf forest, Wa—water, CR—croplands, DB—deciduous broadleaf forest, MF—mixed forest for [CO₂] towers are listed in Table 3. A mixed type is listed if the difference between the fractions of the top two biome types is within 10%.

estimation of the daytime net sink (Figure 8c): decreasing E_{\max} values leads to larger biases across biomes (sites), and increasing E_{\max} values do not offset the weak net uptake in CASA, indicating that tuning the parameters cannot fully counteract the lack of the harvest sink in CASA for summer 2016. The nighttime bias of CASA's downscaled respiration has a smaller bias than the daytime productivity plus respiration, bias is still large, especially when considered as a portion of the averaged nighttime flux magnitude. On average, CASA underestimates the daily averaged sink by $1\text{--}2 \mu \text{mol m}^{-2} \text{s}^{-1}$ for EN, CR, and DB and by $1 \mu \text{mol m}^{-2} \text{s}^{-1}$ or less for GL. The different magnitudes of flux biases across the biomes suggest that the flux bias is scaled with the strength of the seasonal uptake and/or the strength of the annual net carbon exchange (i.e., missing sink).

CR and DB have more distinct E_{\max} ranking patterns for the daytime and nighttime flux biases. For CR, low E_{\max} leads to larger biases in both daytime and nighttime fluxes due to low daytime net uptake and nighttime respiration. Since E_{\max} in CASA directly impacts GPP and indirectly impacts respiration, increasing both GPP and respiration is favorable for capturing CR fluxes in CASA. This is also reflected by the combination of high E_{\max} and high Q_{10} values (P25 and P27) being ranked in the top group in Figure 6. For DB, however, low E_{\max} leads to large, positive flux biases in daytime but the smallest positive bias at night, suggesting that the bigger issue for DB is that the CASA model tends to respire carbon faster than in the actual ecosystem.

In both mole fraction and flux space, CT2017 agrees more closely overall with the observations than the CASA members after averaging (Figure 8). However, once we break nighttime and daytime flux apart, CT2017 overestimates the nighttime respiration for these four biome types and daytime net uptake for DB.

5. Discussion

Multiple CO₂ observation platforms (i.e., aircraft and tower [CO₂] and flux tower data) across measurement space (i.e., concentration vs. flux) identify that E_{\max} is the most important parameter driving the spread of the CASA flux ensemble for summer 2016. In concentration space, the modeled [CO₂] biases can be contributed by biogenic flux, transport, boundary conditions and other CO₂ flux components. Although Figure 3 demonstrates that, in the sampled cases, the biosphere dominates the model errors, the transport errors are not negligible. The model transport errors can increase or decrease modeled [CO₂] biases in Figure 5 as a whole. We choose to focus on the relative performance among these flux members to minimize the impact of the potential transport errors on interpreting the results. However, we acknowledge that the complexity the model transport introduces in the results are not eliminated.

We argue that the transport error is unlikely to change our major conclusion that the biogenic fluxes are biased or change the relative performance across the biogenic CO₂ flux members. This argument is supported in the consistent results from the flux comparison in Section 4.3. This is also reflected by the results that CT2017 has smaller biases against measured [CO₂] (Figures 5 and 6). While the CT2017 fluxes were optimized with the TM5 transport model (Krol et al., 2005), we consistently found that the associated WRF-simulated [CO₂] bias was the lowest compared to the CASA members. We also note that a slightly different version of CASA (Potter et al., 1993) is used as a prior in CT2017. In our transport model setup, we nudged the WRF-Chem model to the state-of-the-art reanalysis product ERA5 in order to improve the large-scale dynamics. The comparison between the nudged and free run modeled wind fields illustrates the effectiveness and efficiency of this meteorological constraint. Finally, the ensemble of perturbed transport simulations confirms the secondary role of transport errors in [CO₂] model-data residuals. Hence, it seems highly unlikely that transport errors could be the cause of our findings regarding the CASA ensemble.

The CarbonTracker (CT) inversion system works by estimating scaling factors that multiply the prior model (CASA)'s NEE. During the growing season, an increase in the terrestrial sink is equivalent to amplifying the prior model's diel cycle, to perhaps unrealistic levels. These scaling factors are constant for each week and are independent from one week to the next. Thus, the CT fluxes may yield periods of unrealistically large diel cycles, possibly followed by periods with abnormally small daily cycles of NEE. The shape of the day-night flux differences and day-to-day changes cannot be adjusted by CT. To alleviate this problem, some regional inverse flux estimation pioneers have applied different strategies. Gourdjji et al. (2012) first estimated diel cycle of the fluxes at a 3-hourly resolution for the North American continent. In Schuh et al. (2013), the two regional inversion techniques were used to minimize the artifact in the diel cycle of the estimated fluxes: the Colorado State University inversion technique optimized for GPP and respiration separately, while the Pennsylvania State University inversion optimized daytime NEE and nighttime NEE on a 7.5-day time scale. Most inversion systems that optimize NEE with prescribed diel cycle may potentially alleviate the unrealistic diel cycles but, on the other side, may be subject to temporal aggregation errors (e.g., Byrne et al., 2020; Peters et al., 2007; Schuh et al., 2010). Additionally, since CT relies heavily on simulated transport, it is also possible to retrieve exaggerated diurnal cycles in fluxes if that model ventilates the PBL too strongly. While this process could affect fluxes throughout the day, thin and stable nighttime boundary layers are particularly difficult to represent in models of this class.

Section 4.3 provides more insight into the causes of the modeled [CO₂] biases for summer 2016. The biggest assumption we made in Equation 1 is homogeneity across the entire biome. The optimal method for calculating the flux biases at a [CO₂] tower should be based on the flux towers in a [CO₂] tower's footprint (shown in Figure S2). However, due to the uneven distribution of the flux towers in the domain of interest, this direct calculation is not possible. The method we used to derive the flux biases at each [CO₂] tower takes the influence from multiple biomes into account and results in a bias ranking among the flux members that is consistent with the ranking in mole fraction space, indicating the reliability of this method.

The CASA simulations were spun up equilibrium, and the parameters were adjusted to match the observations. CASA's overestimation of summer NEE likely derives from missing carbon sink processes

because of the balanced-biosphere equilibrium starting condition resulting in carbon pools that are too large rather than a problem in the model's parameters (Pietsch & Hasenauer, 2006; Wutzler & Reichstein, 2007; Zhou, Williams, Lauvaux, Davis, et al., 2020). The range of the three perturbed parameters, E_{\max} , T_{opt} , and Q_{10} , were determined by comparison to flux measurements over 13-year simulation period. Tuning these parameters cannot solve the fact that carbon stocks are actually dynamic and out of equilibrium. Our results may erroneously indicate that a particular parameter set has the smallest bias when the core problem may be that the carbon pool is out of equilibrium. This is consistent with the finding that CT2017, with its imposed net global land sink, is less biased than the balanced biosphere CASA L2 ensemble. We therefore do not intend to stress which parameter set is the best. Instead, the highlight of this work is to demonstrate that atmospheric $[\text{CO}_2]$ data can be used, with flux tower measurements, to diagnose the performance of TBMs. This study demonstrates the utility of multiple observation platforms for TBM evaluation.

The rankings in Section 4.2 show that increasing E_{\max} does not necessarily lead to better agreement with the measurements. Increasing E_{\max} increases GPP, but the carbon seems to respire away quickly in CASA and does not reside long enough, perhaps because harvest is not represented in this simulation. Key missing processes include the effects of management and land use such as (a) agricultural sinks from management that removes crops and crop residues thus decreasing the size of the carbon pool that might be respired, (b) pastureland sinks from cattle and other grazers that consume plant biomass and store it in their body mass, (c) forest carbon storage as trees and stands mature with sequestration in wood. Additional candidates include stimulation of ecosystem carbon sinks by growth enhancement factors such as rising CO_2 concentration, nitrogen deposition, fertilizer additions, and the like. Including these missing processes in the model framework would require a re-evaluation of all model parameters, and we might find that a different E_{\max} parameter set was best relative to what was identified in this study and might lead to a better match with this suite of observation. In addition, the temporal downscaling of CASA's monthly fluxes suggests that day-night variation in respiration may not be correctly captured in the temporal downscaling to hourly variations. The modeled respiration is primarily related to temperature. However, previous studies have shown that the correlation plots of respiration and temperature are scattered. The high-biased points tend to occur in daytime, and the low-biased points tend to occur at nighttime because carbon fixation processes are more active during daytime when more labile carbon is readily available to be respired.

Simplistic temporal downscaling of CASA's monthly carbon fluxes likely contributes to biases in the diel cycle at sub-daily scale. The 3-hourly CASA ensemble flux products were downscaled from the native monthly resolution using the Olsen and Randerson (2004) method, in which GPP is downscaled with downward shortwave radiation, R_e is downscaled with air temperature and Q_{10} , and NEE is the sum of GPP and R_e . This makes it difficult to use the diel cycle analyses in Section 4.3 to make conclusions about GPP and R_e in CASA at fine scales. However, redistribution of respiration to shift more to nighttime would appear insufficient to address the flux biases. In the flux diurnal analysis, we found a slight over-estimation of nighttime fluxes (bias was near zero for most ecosystems except for DB) and a significant under-estimation of the daytime uptake. If we transfer respiration from daytime to nighttime to fix the daytime sink under-estimation, the net slightly positive flux bias during nighttime will increase, which is contrary to the expected near-zero correction at night or would exacerbate the overestimation where it already exists (again, mainly DB). Hence, changing the diurnal cycle of respiration is insufficient to address the day/night flux mismatches. Consequently, only an increase in GPP (or an overall decrease in respiration) can fix the bias, requiring model structural corrections to include the missing processes mentioned above. Also, note that the biases in the study period of six weeks are nonetheless on a similar timescale as direct estimation from native monthly CASA GPP and R_e .

The ability to diagnose errors in GPP and respiration will require models that resolve the diel cycle and possibly additional data dedicated to the photosynthetic process. Carbonyl sulfide (COS) has been proposed as an independent proxy for GPP as it diffuses into leaves in a fashion very similar to CO_2 , but in contrast to the latter, is generally not emitted by respiration. Campbell et al. (2017) presented a global, measurement-based estimate of GPP growth during the twentieth century that is based on long-term atmospheric COS records, derived from ice-core, firn and ambient air samples and found that the observation-based COS record is

most consistent with simulations of climate and the carbon cycle. Recently, Spielmann et al. (2019) used concurrent ecosystem-scale flux measurements of CO₂ and COS at four European biomes for a joint constraint on CO₂ flux partitioning. Their results demonstrated the importance of using multiple approaches for constraining present-day GPP due to a systematic underestimation under low light conditions with the classical approaches relying merely on CO₂ fluxes. Other studies have used the δ¹³C of nocturnal whole-ecosystem respiration as a proxy from which to derive carbon isotope discrimination associated with photosynthesis (Alstad et al., 2007; Bowling et al., 2002; Flanagan et al., 1996). Joint atmospheric gas constraints may improve our diagnoses of the causes of biases in TBM flux estimates. Studies using ACT-America airborne biogenic tracers, including CO₂, COS, and CO, are submitted to this collection (Parazoo et al., 2020).

6. Conclusions

We evaluate the modeled [CO₂] biases associated with CASA TMB biogenic CO₂ flux ensemble members and the CT2017 posterior biogenic flux using aircraft and tower in situ [CO₂] jointly with eddy-covariance flux data from July 18 to August 28 in summer 2016. Aircraft and tower in situ [CO₂] were influenced by GL, EN, DB, and CR. While the mole fraction-based analyses revealed a systematic underestimation of carbon uptake by the balanced-biosphere CASA runs, the flux-based analyses identified a combined effect from an overestimation of respiration and under-estimation of productivity. The joint observational analysis yields strong confidence that these results span large spatial domains and multiple ecosystems due to the availability of long aircraft transects and a wide network of ground-based measurements. The results from analyzing both mole fraction and flux model-data residuals were consistent. The systematic errors in CASA that span all parameter values suggest that missing processes cannot be properly simulated by adjusting the existing parameters. Analyses indicate that modeled [CO₂] biases are related to biome productivity; the models tend to be biased more for high productivity biomes (DB and CR) and biased less for GL and EN. In particular, the summer harvest sink absent from CASA seems to be responsible for large biases in the Midwest. Lastly, CT2017, an inversion product that is constrained by atmospheric [CO₂] data and has an imposed net biogenic carbon sink, shows better agreement with [CO₂] mole fraction data compared to CASA flux ensemble members. However, the flux-based analyses revealed that the diurnal variations of CT were unrealistically large. We suggest that the scaling of the net daily fluxes in large-scale inversions must be further decomposed into day and night sub-components to reproduce the diel cycle of photosynthetic and ecosystem respiration processes. Lastly, our setup does not resolve the subgrid biome variability, which is a shortcoming at the scale. The CASA ensemble provides an alternative flux data set at 500 × 500 m. A sensitivity study with both sets of the CASA flux ensemble can be tested in future.

Acknowledgments

This work was funded by the NASA ACT-America project. ACT-America project is a NASA Earth Venture Sub-orbital-2 project funded by NASA's Earth Science Division (Grant NNX15AG76G to Penn State and NNX16AN17G to Clark University). S. Feng was also partly supported by the Department of Energy (DOE)'s Energy Esascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The Pacific Northwest National Laboratory is operated by Battelle Memorial Institute under contract DE-AC05-76RL01830. T. Lauvaux was also supported by the French research program Make Our Planet Great Again (project CIUDAD). Many thanks to Andrew R. Jacobson at NOAA Global Monitoring Laboratory (GML) for the feedback on the diel cycle of the CT2017 posterior biogenic CO₂ fluxes.

Data Availability Statement

The CarbonTracker CT2017 product can be obtained from NOAA GML, Boulder, Colorado, USA from the website at <http://carbontracker.noaa.gov>. The L2 CASA ensemble suite used in this work can be found at <https://doi.org/10.3334/ORNLDAAAC/1675>. NCEP Real-Time 12-km SST are archived at <ftp://polar.ncep.noaa.gov/pub/history/sst/ophi>. The ensemble output will be made available from datacommons@psu at <https://doi.org/10.26208/z864-qk7>. The Obspack CO₂ Global View Plus data package can be obtained from <https://www.esrl.noaa.gov/gmd/ccgg/obspack/>. The citations for individual [CO₂] towers are listed in Table S3. AmeriFlux data are obtained from <https://ameriflux.lbl.gov/>, with citations given in Table S2.

References

- Alstad, K. P., Lai, C.-T., Flanagan, L. B., & Ehleringer, J. R. (2007). Environmental controls on the carbon isotope composition of ecosystem-respired CO₂ in contrasting forest ecosystems in Canada and the USA. *Tree Physiology*, 27(10), 1361–1374. <https://doi.org/10.1093/treephys/27.10.1361>
- Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., et al. (2014). CO₂, CO, and CH₄ measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts. *Atmospheric Measurement Techniques*, 7(2), 647–687. <https://doi.org/10.5194/amt-7-647-2014>
- Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P., Denning, A. S., et al. (2006). TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO₂ fluxes, 1988–2003. *Global Biogeochemical Cycles*, 20(1). <https://doi.org/10.1029/2004GB002439>

- Barker, D. M., Huang, W., Guo, Y.-R., Bourgeois, A. J., & Xiao, Q. N. (2004). A three-dimensional variational data assimilation system for MM5: Implementation and initial results. *Monthly Weather Review*, *132*(4), 897–914. [https://doi.org/10.1175/1520-0493\(2004\)132<0897:ATVDAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0897:ATVDAS>2.0.CO;2)
- Basu, S., Miller, J. B., & Lehman, S. (2016). Separation of biospheric and fossil fuel fluxes of CO₂ by atmospheric inversion of CO₂ and ¹⁴C measurements: Observation system simulations. *Atmospheric Chemistry and Physics*, *16*(9), 5665–5683. <https://doi.org/10.5194/acp-16-5665-2016>
- Berner, J., Fossell, K. R., Ha, S.-Y., Hacker, J. P., & Snyder, C. (2015). Increasing the skill of probabilistic forecasts: Understanding performance improvements from model-error representations. *Monthly Weather Review*, *143*(4), 1295–1320. <https://doi.org/10.1175/MWR-D-14-00091.1>
- Berner, J., Ha, S.-Y., Hacker, J. P., Fournier, A., & Snyder, C. (2011). Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Monthly Weather Review*, *139*(6), 1972–1995. <https://doi.org/10.1175/2010MWR3595.1>
- Berner, J., Shutts, G. J., Leutbecher, M., & Palmer, T. N. (2009). A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *Journal of the Atmospheric Sciences*, *66*(3), 603–626. <https://doi.org/10.1175/2008JAS2677.1>
- Bowling, D. R., McDowell, N. G., Bond, B. J., Law, B. E., & Ehleringer, J. R. (2002). 13C content of ecosystem respiration is linked to precipitation and vapor pressure deficit. *Oecologia*, *131*(1), 113–124. <https://doi.org/10.1007/s00442-001-0851-y>
- Butler, M. P., Lauvaux, T., Feng, S., Liu, J., Bowman, K. W., & Davis, K. J. (2020). Atmospheric simulations of total column CO₂ Mole fractions from global to mesoscale within the carbon monitoring system flux inversion framework. *Atmosphere*, *11*(8), 787. <https://doi.org/10.3390/atmos11080787>
- Byrne, B., Liu, J., Lee, M., Baker, I., Bowman, K. W., Deutscher, N. M., et al. (2020). Improved constraints on northern extratropical CO₂ fluxes obtained by combining surface-based and space-based atmospheric CO₂ measurements. *Journal of Geophysical Research: Atmospheres*, *125*(15), e2019JD032029. <https://doi.org/10.1029/2019JD032029>
- Campbell, J. E., Berry, J. A., Seibt, U., Smith, S. J., Montzka, S. A., Launois, T., et al. (2017). Large historical growth in global terrestrial gross primary production. *Nature*, *544*(7648), 84–87. <https://doi.org/10.1038/nature22030>
- Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 569–585. [https://doi.org/10.1175/1520-0493\(2001\)129<0569:CAALSH>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2)
- Cooperative Global Atmospheric Data Integration Project. (2019). *Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957–2018; obspack_co2_1_GLOBALVIEWplus_v5.0_2019_08_12*. 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>
- 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*, *1*(aop), 1–54. <https://doi.org/10.1175/BAMS-D-20-0300.1>
- Davis, K. J., Obland, M. D., Lin, B., Lauvaux, T., O'dell, C., Meadows, B., et al. (2018). ACT-America: L3 merged in situ atmospheric trace gases and flask data, Eastern USA. ORNL DAAC. <https://doi.org/10.3334/ORNLDAAC/1593>
- Diaz-Isaac, L. I., Lauvaux, T., & Davis, K. J. (2018). Impact of physical parameterizations and initial conditions on simulated atmospheric transport and CO₂ mole fractions in the US Midwest. *Atmospheric Chemistry and Physics*, *18*(20), 14813–14835. <https://doi.org/10.5194/acp-18-14813-2018>
- Enting, I. G., Trudinger, C. M., & Francey, R. J. (1995). A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO₂. *Tellus B: Chemical and Physical Meteorology*, *47*(1–2), 35–52. <https://doi.org/10.1034/j.1600-0889.47.issue1.5.x>
- Feng, S., Lauvaux, T., Davis, K. J., Keller, K., Zhou, Y., Williams, C., et al. (2019). Seasonal characteristics of model uncertainties from biogenic fluxes, transport, and large-scale boundary inflow in atmospheric CO₂ simulations over North America. *Journal of Geophysical Research: Atmospheres*, *124*(24), 14325–14346. <https://doi.org/10.1029/2019JD031165>
- Feng, S., Lauvaux, T., Keller, K., Davis, K. J., Rayner, P., Oda, T., & Gurney, K. R. (2019). A road map for improving the treatment of uncertainties in high-resolution regional carbon flux inverse estimates. *Geophysical Research Letters*, *46*(22), 13461–13469. <https://doi.org/10.1029/2019GL082987>
- Flanagan, L. B., Brooks, J. R., Varney, G. T., Berry, S. C., & Ehleringer, J. R. (1996). Carbon isotope discrimination during photosynthesis and the isotope ratio of respired CO₂ in boreal forest ecosystems. *Global Biogeochemical Cycles*, *10*(4), 629–640. <https://doi.org/10.1029/96GB02345>
- Friedlingstein, P., Jones, M. W., O'sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., et al. (2019). Global carbon budget 2019. *Earth System Science Data*, *11*(4), 1783–1838. <https://doi.org/10.5194/essd-11-1783-2019>
- Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, *27*(2), 511–526. <https://doi.org/10.1175/JCLI-D-12-00579.1>
- Fung, I. Y., Tucker, C. J., & Prentice, K. C. (1987). Application of Advanced Very High Resolution Radiometer vegetation index to study atmosphere-biosphere exchange of CO₂. *Journal of Geophysical Research*, *92*(D3), 2999–3015. <https://doi.org/10.1029/JD092iD03p02999>
- Gloor, M., Bakwin, P., Hurst, D., Lock, L., Draxler, R., & Tans, P. (2001). What is the concentration footprint of a tall tower? *Journal of Geophysical Research*, *106*(D16), 17831–17840. <https://doi.org/10.1029/2001JD900021>
- Gonçalves, L. G. G. D., Borak, J. S., Costa, M. H., Saleska, S. R., Baker, I., Restrepo-Coupe, N., et al. (2013). Overview of the large-scale biosphere-atmosphere experiment in Amazonia data model intercomparison project (LBA-DMIP). *Agricultural and Forest Meteorology*, *182–183*, 111–127. <https://doi.org/10.1016/j.agrformet.2013.04.030>
- Gourdji, S. M., Mueller, K. L., Yadav, V., Huntzinger, D. N., Andrews, A. E., Trudeau, M., et al. (2012). North American CO₂ exchange: Inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion. *Biogeosciences*, *9*(1), 457–475. <https://doi.org/10.5194/bg-9-457-2012>
- Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled “online” chemistry within the WRF model. *Atmospheric Environment*, *39*(37), 6957–6975. <https://doi.org/10.1016/j.atmosenv.2005.04.027>
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., et al. (2002). Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models. *Nature*, *415*(6872), 626–630. <https://doi.org/10.1038/415626a>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Huntzinger, D. N., Gourdji, S. M., Mueller, K. L., & Michalak, A. M. (2011). The utility of continuous atmospheric measurements for identifying biospheric CO₂ flux variability. *Journal of Geophysical Research*, *116*(D6). <https://doi.org/10.1029/2010JD015048>

- Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., et al. (2013). The North American carbon program multi-scale synthesis and terrestrial model intercomparison project—Part 1: Overview and experimental design. *Geoscientific Model Development*, 6(6), 2121–2133. <https://doi.org/10.5194/gmd-6-2121-2013>
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113(D13). <https://doi.org/10.1029/2008JD009944>
- Kain, J. S. (2004). The Kain–Fritsch convective parameterization: An update. *Journal of Applied Meteorology*, 43(1), 170–181. [https://doi.org/10.1175/1520-0450\(2004\)043<0170:TKCPAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2)
- Kaminski, T., Knorr, W., Rayner, P. J., & Heimann, M. (2002). Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle. *Global Biogeochemical Cycles*, 16(4), 14–14–16. <https://doi.org/10.1029/2001GB001463>
- Kaminski, T., Knorr, W., Scholze, M., Gobron, N., Pinty, B., Giering, R., & Mathieu, P.-P. (2012). Consistent assimilation of MERIS FAPAR and atmospheric CO₂ into a terrestrial vegetation model and interactive mission benefit analysis. *Biogeosciences*, 9(8), 3173–3184. <https://doi.org/10.5194/bg-9-3173-2012>
- Knorr, W. (2000). Annual and interannual CO₂ exchanges of the terrestrial biosphere: Process-based simulations and uncertainties. *Global Ecology and Biogeography*, 9(3), 225–252. <https://doi.org/10.1046/j.1365-2699.2000.00159.x>
- Koffi, E. N., Rayner, P. J., Scholze, M., Chevallier, F., & Kaminski, T. (2013). Quantifying the constraint of biospheric process parameters by CO₂ concentration and flux measurement networks through a carbon cycle data assimilation system. *Atmospheric Chemistry and Physics*, 13(21), 10555–10572. <https://doi.org/10.5194/acp-13-10555-2013>
- Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., et al. (2005). The two-way nested global chemistry-transport zoom model TM5: Algorithm and applications. *Atmospheric Chemistry and Physics*, 5(2), 417–432. <https://doi.org/10.5194/acp-5-417-2005>
- Lauvaux, T., Schuh, A. E., Uliasz, M., Richardson, S., Miles, N., Andrews, A. E., et al. (2012). Constraining the CO₂ budget of the corn belt: Exploring uncertainties from the assumptions in a mesoscale inverse system. *Atmospheric Chemistry and Physics*, 12(1), 337–354. <https://doi.org/10.5194/acp-12-337-2012>
- Liu, J., Bowman, K. W., Lee, M., Henze, D. K., Bousserez, N., Brix, H., et al. (2014). Carbon monitoring system flux estimation and attribution: Impact of ACOS-GOSAT XCO₂ sampling on the inference of terrestrial biospheric sources and sinks. *Tellus B: Chemical and Physical Meteorology*, 66(1), 22486. <https://doi.org/10.3402/tellusb.v66.22486>
- Luo, Y. Q., Randerson, J. T., Abramowitz, G., Bacour, C., Blyth, E., Carvalhais, N., et al. (2012). A framework for benchmarking land models. *Biogeosciences*, 9(10), 3857–3874. <https://doi.org/10.5194/bg-9-3857-2012>
- Masarie, K. A., Peters, W., Jacobson, A. R., & Tans, P. P. (2014). ObsPack: A framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements. *Earth System Science Data*, 6(2), 375–384. <https://doi.org/10.5194/essd-6-375-2014>
- McCaughy, J. H., Pejam, M. R., Arain, M. A., & Cameron, D. A. (2006). Carbon dioxide and energy fluxes from a boreal mixedwood forest ecosystem in Ontario, Canada. *Agricultural and Forest Meteorology*, 140(1), 79–96. <https://doi.org/10.1016/j.agrformet.2006.08.010>
- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North American regional reanalysis. *Bulletin of the American Meteorological Society*, 87(3), 343–360. <https://doi.org/10.1175/bams-87-3-343>
- Nakanishi, M., & Niino, H. (2006). An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*, 119(2), 397–407. <https://doi.org/10.1007/s10546-005-9030-8>
- Ogle, S. M., Davis, K., Lauvaux, T., Schuh, A., Cooley, D., West, T. O., et al. (2015). An approach for verifying biogenic greenhouse gas emissions inventories with atmospheric CO₂ concentration data. *Environmental Research Letters*, 10(3), 034012. <https://doi.org/10.1088/1748-9326/10/3/034012>
- Olsen, S. C., & Randerson, J. T. (2004). Differences between surface and column atmospheric CO₂ and implications for carbon cycle research. *Journal of Geophysical Research*, 109(D2). <https://doi.org/10.1029/2003JD003968>
- 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(5). <https://doi.org/10.1029/2019jd030526>
- Parazoo, N. C., Bowman, K. W., Baier, B. C., Liu, J., Lee, M., Kuai, L., et al. (2020). December 28Covariation of airborne biogenic tracers (CO₂, COS, and CO) supports stronger than expected growing season photosynthetic uptake in the southeastern US. (preprint). <https://doi.org/10.1002/essoar.10505574.1>
- Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., et al. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences of the United States of America*, 104(48), 18925–18930. <https://doi.org/10.1073/pnas.0708986104>
- 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 CO₂ fluxes estimated from OCO-2 retrievals. *Atmospheric Chemistry and Physics*, 19(20), 13267–13287. <https://doi.org/10.5194/acp-19-13267-2019>
- Pietsch, S. A., & Hasenauer, H. (2006). Evaluating the self-initialization procedure for large-scale ecosystem models. *Global Change Biology*, 12(9), 1658–1669. <https://doi.org/10.1111/j.1365-2486.2006.01211.x>
- Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B., & Heimann, M. (2012). Comparing Lagrangian and Eulerian models for CO₂ transport—A step towards Bayesian inverse modeling using WRF/STILT-VPRM. *Atmospheric Chemistry and Physics*, 12(19), 8979–8991. <https://doi.org/10.5194/acp-12-8979-2012>
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, 7(4), 811–841. <https://doi.org/10.1029/93GB02725>
- Samaddar, A., Feng, S., Lauvaux, T., Barkley, Z. R., Pal, S., & Davis, K. J. (2021). Carbon dioxide distribution, origins, and transport along a frontal boundary during summer in mid-latitudes. *Journal of Geophysical Research: Atmospheres*, e2020JD033118. <https://doi.org/10.1029/2020JD033118>
- Sarrat, C., Noilhan, J., Dolman, A. J., Gerbig, C., Ahmadov, R., Tolk, L. F., et al. (2007). Atmospheric CO₂ modeling at the regional scale: An intercomparison of 5 meso-scale atmospheric models. *Biogeosciences*, 4(6), 1115–1126. <https://doi.org/10.5194/bg-4-1115-2007>
- Scholze, M., Kaminski, T., Knorr, W., Voßbeck, M., Wu, M., Ferrazzoli, P., et al. (2019). Mean European Carbon sink over 2010–2015 estimated by simultaneous assimilation of atmospheric CO₂, soil moisture, and vegetation optical depth. *Geophysical Research Letters*, 46(23), 13796–13803. <https://doi.org/10.1029/2019GL085725>
- Schuh, A. E., Denning, A. S., Corbin, K. D., Baker, I. T., Uliasz, M., Parazoo, N., et al. (2010). A regional high-resolution carbon flux inversion of North America for 2004. *Biogeosciences*, 7(5), 1625–1644. <https://doi.org/10.5194/bg-7-1625-2010>

- 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 CO₂ surface flux estimates. *Global Biogeochemical Cycles*, 33(4), 484–500. <https://doi.org/10.1029/2018GB006086>
- Schuh, A. E., Lauvaux, T., West, T. O., Denning, A. S., Davis, K. J., Miles, N., et al. (2013). Evaluating atmospheric CO₂ inversions at multiple scales over a highly inventoried agricultural landscape. *Global Change Biology*, 19(5), 1424–1439. <https://doi.org/10.1111/gcb.12141>
- Schwalm, C. R., Huntzinger, D. N., Fisher, J. B., Michalak, A. M., Bowman, K., Ciais, P., et al. (2015). Toward “optimal” integration of terrestrial biosphere models. *Geophysical Research Letters*, 42(11), 4418–4428. <https://doi.org/10.1002/2015GL064002>
- Seibert, P., & Frank, A. (2004). Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. *Atmospheric Chemistry and Physics*, 4(1), 51–63. <https://doi.org/10.5194/acp-4-51-2004>
- Shutts, G. (2005). A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 3079–3102. <https://doi.org/10.1256/qj.04.106>
- Sitch, S., Huntingford, C., Gedney, N., Levy, P. E., Lomas, M., Piao, S. L., et al. (2008). Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, 14(9), 2015–2039. <https://doi.org/10.1111/j.1365-2486.2008.01626.x>
- Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Barker, D., Duda, G., et al. (2008). *A description of the advanced research WRF version 3*. <https://doi.org/10.5065/D68S4MVH>
- Spielmann, F. M., Wohlfahrt, G., Hammerle, A., Kitz, F., Migliavacca, M., Alberti, G., et al. (2019). Gross primary productivity of four European ecosystems constrained by joint CO₂ and COS flux measurements. *Geophysical Research Letters*, 46(10), 5284–5293. <https://doi.org/10.1029/2019GL082006>
- Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., et al. (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. <https://doi.org/10.1002/2014JD022591>
- Tans, P. P., & Conway, T. J. (2005). *Monthly atmospheric CO₂ mixing ratios from the NOAA CMDL carbon cycle cooperative global air sampling network, 1968-2002 (No. cdiaac:NOAA/CMDL)*. Environmental System Science Data Infrastructure for a Virtual Ecosystem. <https://doi.org/10.15485/1464220>
- Tans, P. P., Fung, I. Y., & Takahashi, T. (1990). Observational constraints on the global atmospheric CO₂ budget. *Science*, 247(4949), 1431–1438. <https://doi.org/10.1126/science.247.4949.1431>
- Thompson, G., Rasmussen, R. M., & Manning, K. (2004). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Monthly Weather Review*, 132(2), 519–542. [https://doi.org/10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2)
- Uliasz, M., Bartochowska, M., Madany, A., Piwkowski, H., Parfianiewicz, J., & Rozkrut, M. (1994). Application of the mesoscale dispersion modeling system to investigation of air pollution transport in Southern Poland. In S.-E. Gryning, & M. M. Millán (Eds.), *Air pollution modeling and its application X* (pp. 27–34). Springer US. https://doi.org/10.1007/978-1-4615-1817-4_4
- United Nations Framework Convention on Climate Change (UNFCCC). (2015). *Decision 1/CP.21: Adoption of the Paris agreement (No. FCCC/CP/2015/L.9/Rev.1. 12 Dec)*. UNFCCC. Retrieved from <https://unfccc.int/resource/docs/2015/cop21/eng/109r01.pdf>
- Wutzler, T., & Reichstein, M. (2007). Soils apart from equilibrium—Consequences for soil carbon balance modelling. *Biogeosciences*, 4(1), 125–136. <https://doi.org/10.5194/bg-4-125-2007>
- Zhou, Y., Williams, C. A., Lauvaux, T., Davis, K. J., Feng, S., Baker, I., et al. (2020). A Multiyear gridded data ensemble of surface biogenic carbon fluxes for North America: Evaluation and analysis of results. *Journal of Geophysical Research: Biogeosciences*, 125(2), e2019JG005314. <https://doi.org/10.1029/2019JG005314>
- Zhou, Y., Williams, C. A., Lauvaux, T., Feng, S., Baker, I. T., Wei, Y., et al. (2020). *ACT-America: Gridded ensembles of surface biogenic carbon fluxes, 2003-2019*. <https://doi.org/10.3334/ORNLDAAAC/1675>

References From the Supporting Information

- McClure, C. D., Jaffe, D. A., & Gao, H. (2016). Carbon dioxide in the free troposphere and boundary layer at the Mt. Bachelor Observatory. *Aerosol and Air Quality Research*, 16(3), 717–728. <https://doi.org/10.4209/aaqr.2015.05.0323>
- Miles, N. L., Martins, D. K., Richardson, S. J., Rella, C. W., Arata, C., Lauvaux, T., et al. (2018). Calibration and field testing of cavity ring-down laser spectrometers measuring CH₄, CO₂, and δ¹³CH₄ deployed on towers in the Marcellus Shale region. *Atmospheric Measurement Techniques*, 11(3), 1273–1295. <https://doi.org/10.5194/amt-11-1273-2018>
- Miles, N. L., Richardson, S. J., Martins, D. K., Davis, K. J., Lauvaux, T., Haupt, B. J., & Miller, S. K. (2018). *ACT-America: L2 in situ CO₂, CO, and CH₄ Concentrations from Towers, Eastern USA*. <https://doi.org/10.3334/ORNLDAAAC/1568>
- Sargent, M., Barrera, Y., Nehrkorn, T., Huttyra, L. R., Gately, C. K., Jones, T., et al. (2018). Anthropogenic and biogenic CO₂ fluxes in the Boston urban region. *Proceedings of the National Academy of Sciences of the United States of America*, 115(29), 7491–7496. <https://doi.org/10.1073/pnas.1803715115>
- Schmidt, A., Law, B. E., Göckede, M., Hanson, C., Yang, Z., & Conley, S. (2016). Bayesian optimization of the Community Land Model simulated biosphere-atmosphere exchange using CO₂ observations from a dense tower network and aircraft campaigns over Oregon. *Earth Interactions*, 20, 1–35. <https://doi.org/10.1175/EI-D-16-0011.1>
- Worthy, D. E. J., Higuchi, K., & Chan, D. (2003). North American influence on atmospheric carbon dioxide data collected at Sable Island, Canada. *Tellus B: Chemical and Physical Meteorology*, 55(2), 105–114. <https://doi.org/10.3402/tellusb.v55i2.16731>