Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2020GB006914

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

Sha Feng1,2, Thomas Lauvaux1,3, Christopher A. Williams4, Kenneth J. Davis1,4, Yu Zhou4, Ian Baker4, Zachary R. Barkley4, 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 (TBM) 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 $E_{\text{max}}$ 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 atmospheric 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 biologic 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 CO$_2$ 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 [CO$_2$] from towers or satellites (e.g., Scholze et al., 2019). On sub-continental scales, atmospheric [CO$_2$] 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 CO$_2$ 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 [CO$_2$] 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 [CO$_2$] errors attributed to biogenic CO$_2$ fluxes, fossil fuel emissions, atmospheric transport, and large-scale boundary inflow from daily to annual timescales and discovered that the biogenic CO$_2$ fluxes dominate the model errors across timescales, implying that [CO$_2$] observations hold promise for evaluating and improving TBMs. Global inversions usually optimize CO$_2$ 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 CO$_2$ 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 CO$_2$ 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 $E_{\text{max}}$, optimal temperature of photosynthesis $T_{\text{opt}}$, and maximum light use efficiency $E_{\text{max}}$. They found that the diurnal variability in surface fluxes within the near field of tower [CO$_2$] 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 [CO$_2$] errors attributed to biogenic CO$_2$ fluxes, fossil fuel emissions, atmospheric transport, and large-scale boundary inflow from daily to annual timescales and discovered that the biogenic CO$_2$ fluxes dominate the model errors across timescales, implying that [CO$_2$] observations hold promise for evaluating and improving TBMs. Global inversions usually optimize CO$_2$ 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 CO$_2$ 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 CO$_2$ 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 $E_{\text{max}}$, optimal temperature of photosynthesis $T_{\text{opt}}$, and maximum light use efficiency $E_{\text{max}}$. They found that the diurnal variability in surface fluxes within the near field of tower [CO$_2$] 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 [CO$_2$] errors attributed to biogenic CO$_2$ fluxes, fossil fuel emissions, atmospheric transport, and large-scale boundary inflow from daily to annual timescales and discovered that the biogenic CO$_2$ fluxes dominate the model errors across timescales, implying that [CO$_2$] observations hold promise for evaluating and improving TBMs. Global inversions usually optimize CO$_2$ 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 CO$_2$ 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 CO$_2$ 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 $E_{\text{max}}$, optimal temperature of photosynthesis $T_{\text{opt}}$, and maximum light use efficiency $E_{\text{max}}$. They found that the diurnal variability in surface fluxes within the near field of tower [CO$_2$] 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 [CO$_2$] errors attributed to biogenic CO$_2$ fluxes, fossil fuel emissions, atmospheric transport, and large-scale boundary inflow from daily to annual timescales and discovered that the biogenic CO$_2$ fluxes dominate the model errors across timescales, implying that [CO$_2$] observations hold promise for evaluating and improving TBMs. Global inversions usually optimize CO$_2$ 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 CO$_2$ 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.
$T_{op}$, 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 [CO$_2$] measurements provide another opportunity for evaluating modeled biogenic CO$_2$ fluxes, as evidenced by the fact that continental [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 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 [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 [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 ([CO$_2$]bio) hereafter) can be directly calculated by convolving biogenic CO$_2$ fluxes with the influencing areas of the [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 [CO$_2$]bio due to mixed signals in the [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 [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 [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 [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 [CO$_2$] exist across frontal boundaries based on the data collected from the summer 2016 campaign. The cross-frontal [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 [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 [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 [CO$_2$]
measurements and tower measurements yield consistent evaluations of the modeled fluxes. Merging these two \([\text{CO}_2]\) 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 \([\text{CO}_2]\) simulated using the CASA L2 flux ensemble members to the airborne and tower-based \([\text{CO}_2]\) 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 \([\text{CO}_2]\), 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 \([\text{CO}_2]\) fluxes (Section 2.2) and \([\text{CO}_2]\) boundary conditions (Section 2.4) in one transport run (Section 2.3). In each transport run, WRF-Chem carries 39 \([\text{CO}_2]\) tracers for the ensembles of boundary conditions and biogenic fluxes, ocean, fossil fuel, and biomass burning fluxes. The modeled total \([\text{CO}_2]\) is the sum of a boundary condition, a biogenic flux, and the oceanic flux, fossil fuel emission, and biomass burning \([\text{CO}_2]\) tracers as described in Feng, Lauvaux, Davies, et al. (2019) and Feng, Lauvaux, Keller, et al. (2019). The \([\text{CO}_2]\) boundary condition tracers are propagated into WRF-Chem hourly with the consideration of the conservation of mass (Butler et al., 2020). Five global \([\text{CO}_2]\) inversion/reanalysis systems are used for \([\text{CO}_2]\) boundary conditions (Section 2.4) and 29 biogenic \([\text{CO}_2]\) fluxes are used for the biogenic \([\text{CO}_2]\) tracers (Section 2.2). The \([\text{CO}_2]\) 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 \(27 \times 27\) 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 \([\text{CO}_2]\) Flux Ensemble: CASA and CT2017

We include 29 biogenic \([\text{CO}_2]\) 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 \([\text{CO}_2]\) flux.
The CASA ensemble members (Zhou, Williams, Lauvaux, Feng, et al., 2020) were generated by perturbing the maximum light use efficiency (\(E_{\text{max}}\)), optimal temperature of photosynthesis (\(T_{\text{opt}}\)), 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, 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 [CO\(_2\)] (Peters et al., 2007). CT2017 global 3-hourly posterior biogenic fluxes at 1° × 1° 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 [CO\(_2\)] and demonstrated that a relatively small transport ensemble can represent model [CO\(_2\)] transport uncertainty. We use the root-mean-square deviation (RMSD) of the simulated [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 Diaz-Issaq 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.

### Table 2

| 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.
2.4. Boundary Condition Ensemble

A suite of optimized [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 [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 [CO$_2$] measurements, we also use a subset of the ground-based [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 [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 [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 [CO$_2$] biases and eliminate the potential interference of transport and boundary conditions inherent in [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 [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 [CO$_2$] tower.

Due to the uneven distribution of flux and [CO$_2$] towers, flux towers cannot be always found in the footprint of a given [CO$_2$] tower. We therefore propagate the flux biases at each [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, $F_\beta$, and bias for the given biome. The flux biases at a given [CO$_2$] tower location $B_n$ can be expressed in the following equation.

$$B_n = \frac{1}{24} \sum_{i=1}^{24} \left( \sum_{j=0}^{14} F_{i,j} W_j \right) W_i,$$

where $W_i$ and $W_j$ are the spatial and temporal weighting functions that propagate the flux biases from the biome level to the [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. $W_i$ are the fractional areas of individual biomes relative to the entire summer-averaged influence area of
Global Biogeochemical Cycles

In the analysis, we find that some towers can be influenced by up to six biome types according to the footprint of the tower \([\mathrm{CO}_2]\) 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, airmass 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 \([\mathrm{CO}_2]\) along the flight tracks.

3.2. NOAA ObsPack GlobalViewPlus [\CO_2] Product

We also use tower-based in situ \([\mathrm{CO}_2]\) data from the NOAA ObsPack GlobalViewPlus product (Cooperative Global Global Atmospheric Data Integration Project, 2019; see locations in Figure 1). Twenty-three tower locations...
Table 3

| 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 Ob-Stack 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.
Global Biogeochemical Cycles

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-
Aligning the aircraft and tower comparisons together, the CASA-simulated \([\text{CO}_2]\) tend to somewhat underestimate the variability of aircraft \([\text{CO}_2]\) (root-mean-squared errors) but overestimate that of tower \([\text{CO}_2]\). 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 \([\text{CO}_2]\) but overestimate temporal variability at the synoptic scale.

Another feature worth noting is that the CASA-simulated \([\text{CO}_2]\) are clustered into two groups. Some high \(E_{\text{max}}\) members fall into the low \(E_{\text{max}}\) group, and the rest falls in the medium \(E_{\text{max}}\) group. This grouping is evident in both aircraft and tower \([\text{CO}_2]\) 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 \([\text{CO}_2]\) 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 \([\text{CO}_2]\) 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 \([\text{CO}_2]\). This is also reflected in Figure 3. Figure 3 also displays that both biogenic flux and transport can play an important role in modeled \([\text{CO}_2]\) 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 \([\text{CO}_2]\) is more positively biased. Assuming that spatial variability in \([\text{CO}_2]\) 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 \([\text{CO}_2]\) data. However, the fact that WRF-Chem performs better when coupled to optimized CT2017 biogenic \(\text{CO}_2\) fluxes confirms that transport model differences remain much smaller than flux differences.
Figure 5. Modeled $[\text{CO}_2]$ 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 $[\text{CO}_2]$ 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 $E_{\text{max}}$ ($E_1$), CASA medium $E_{\text{max}}$ ($E_2$), and CASA high $E_{\text{max}}$ ($E_3$) groups. “P#” on the x-axis are the indexes of CASA ensemble members in the original product.
For the second aspect, we discover that distinct groups of CASA members reflect the three parameter values for maximum light use efficiency, $E_{\text{max}}$. The flux members that show the best agreement with the observations mostly have medium $E_{\text{max}}$ values while the groups with low and high $E_{\text{max}}$ values correspond to larger model biases. Both aircraft and tower measurements identify that $E_{\text{max}}$ 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 $E_{\text{max}}$ ($E_2$) leads to better modeled $[\text{CO}_2]$ across different biome types and observation platforms. The groups with low ($E_1$) and high ($E_3$) $E_{\text{max}}$ tend to have similar biases. A few members with high $E_{\text{max}}$ 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. However, due to the impact of the long-range transport over upwind biomes ($E_2$ not equal $E_3$), the model $[\text{CO}_2]$ bias ranking is different for the DB’s and EN’s $E_{\text{max}}$ groups.

Further investigating the top-performing 11 ensemble members according to $Q_{10}$ values (Figure 7), we find that in general, the flux members with low $Q_{10}$ value ($Q_{10} = 1.2$) are ranked high, followed with medium $Q_{10}$ ($Q_{10} = 1.4$), and then high $Q_{10}$ value ($Q_{10} = 1.6$). In contrast, no $T_{\text{opt}}$-driven grouping is visible (not shown). We conclude that $Q_{10}$ plays a secondary role in CASA-simulated summer NEE.
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$_2$] 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 $E_{\text{max}}$ groups and CT2017 with respect to flux tower data is consistent with the comparison to [CO$_2$]. The low $E_{\text{max}}$ 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 $-1$ μmol m$^{-2}$ s$^{-1}$, from $-8$ to $-3$ μmol m$^{-2}$ s$^{-1}$, and from $0.5$ to $2.5$ μmol m$^{-2}$ s$^{-1}$, respectively (Figure S5 in SI). Temporal downscaling of CASA's monthly fluxes resulted in consistent under-

![Table: Biogenic flux members comparison](image)
estimation of the daytime net sink (Figure 8c): decreasing $E_{\text{max}}$ values leads to larger biases across biomes (sites), and increasing $E_{\text{max}}$ 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 1–2 µmol m$^{-2}$ s$^{-1}$ for EN, CR, and DB and by 1 µmol m$^{-2}$ 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 $E_{\text{max}}$ ranking patterns for the daytime and nighttime flux biases. For CR, low $E_{\text{max}}$ leads to larger biases in both daytime and nighttime fluxes due to low daytime net uptake and nighttime respiration. Since $E_{\text{max}}$ 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 $E_{\text{max}}$ and high $Q_{10}$ values (P25 and P27) being ranked in the top group in Figure 6. For DB, however, low $E_{\text{max}}$ 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 CO\textsubscript{2} observation platforms (i.e., aircraft and tower [CO\textsubscript{2}] and flux tower data) across measurement space (i.e., concentration vs. flux) identify that \(E_{\text{max}}\) is the most important parameter driving the spread of the CASA flux ensemble for summer 2016. In concentration space, the modeled [CO\textsubscript{2}] biases can be contributed by biogenic flux, transport, boundary conditions and other CO\textsubscript{2} 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 [CO\textsubscript{2}] 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 of 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 CO\textsubscript{2} 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 [CO\textsubscript{2}] (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 [CO\textsubscript{2}] 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 [CO\textsubscript{2}] 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. Gouroji 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 [CO\textsubscript{2}] 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 [CO\textsubscript{2}] tower should be based on the flux towers in a [CO\textsubscript{2}] 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 [CO\textsubscript{2}] 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, \( E_{\text{max}} \), \( T_{\text{opt}} \), and \( Q_{\text{in}} \), 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 \( E_{\text{max}} \) does not necessarily lead to better agreement with the measurements. Increasing \( E_{\text{max}} \) 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 \( \text{CO}_2 \) 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 \( E_{\text{max}} \) 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 \( Q_{\text{in}} \), 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 expect 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 mismatch. 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 \( \text{CO}_2 \), 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 CO$_2$ and COS at four European biomes for a joint constraint on CO$_2$ 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 CO$_2$ fluxes. Other studies have used the $\delta^{13}$C 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 CO$_2$, COS, and CO, are submitted to this collection (Parazoo et al., 2020).

6. Conclusions

We evaluate the modeled [CO$_2$] biases associated with CASA TMB biogenic CO$_2$ flux ensemble members and the CT2017 posterior biogenic flux using aircraft and tower in situ [CO$_2$] jointly with eddy-covariance flux data from July 18 to August 28 in summer 2016. Aircraft and tower in situ [CO$_2$] 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 [CO$_2$] 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 [CO$_2$] data and has an imposed net biogenic carbon sink, shows better agreement with [CO$_2$] 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.

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/ORNLDAAC/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 CO$_2$ Global View Plus data package can be obtained from https://www.esrl.noaa.gov/gmd/ccgg/obspack/. The citations for individual [CO$_2$] 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 CO$_2$ 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., Koeller, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., et al. (2014). CO$_2$, CO, and CH$_4$ measurements from tall towers in the NOAA Earth System Research Laboratory’s Global Greenhouse Gas Reference Network: Instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts. Atmospheric Measurement Techniques, 7(2), 647–687. https://doi.org/10.5194/amt-7-647-2014

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 CO$_2$ 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

Basu, S., Miller, J. B., & Lehman, S. (2016). Separation of biospheric and fossil fuel fluxes of CO₂ by atmospheric inversion of CO₂ and 13CO₂ measurements: Observation system simulations. *Atmospheric Chemistry and Physics, 16*(9), 5665–5683. https://doi.org/10.5194/acp-16-5665-2016

Berner, J., Possell, K. R., Ha, S.-Y., Hacker, I. 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 multiparameters representations. *Monthly Weather Review, 139*(6), 1972–1995. https://doi.org/10.1175/2010MWR3195.1

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. F., 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

Cooperative Global Atmospheric Data Integration Project. (2019). Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2018; obspack_co2_v1.0. https://doi.org/10.5194/1299_2019_08_12_NOAA_Earth_System_Research_Laboratory_Global_Monitoring_Division

Crowell, S., Baker, D., Schub, 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, 102*(aap), 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 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 13C 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 in low 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/10.1029/95GB02345

Friedlingstein, P., Jones, M. W., O’Neill, 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., Channan, D. C., 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., Basu, S., Miller, J. B., Lehman, S., & Lehman, S. (2016). Separation of biospheric and fossil fuel fluxes of CO₂ by atmospheric inversion of CO₂ and 13CO₂ measurements: Observation system simulations. *Atmospheric Chemistry and Physics, 16*(9), 5665–5683. https://doi.org/10.5194/acp-16-5665-2016

González, V., G. D., Bézard, A. J., Puech, F., Hiriart, H., Horanyi, A., & 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(1), 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:Tkfcpc>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), 1-14-16. https://doi.org/10.1029/2001GB001483

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/10.1111/j.1466-2528.2000.00159.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

Kolb, 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., Bousher, N., Bris, H., et al. (2014). Carbon monitoring system flux estimation and attribution: Impact of ACO2-GOSSAT XCO2 sampling on the inference of terrestrial biospheric sources and sinks. Tellus B: Chemical and Physical Meteorology, 66(1), 22486. https://doi.org/10.10342/telluvb.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

McCaughey, 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., Shaffar, 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 28 Covariation of airborne biogenic tracers (CO2, COS, and CO) supports stronger than expected growing season photosynthetic uptake in the southeastern US. (preprint). https://doi.org/10.1029/2020GB006914

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.0708983104

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

Pilillo, D., Gerbig, C., Kreieckracher, 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 ecosystems for 2000. (preprint). https://doi.org/10.1016/j.agrformet.2006.08.010

Sarrat, C., Noilhan, J., Dolman, A. J., Gerbig, C., Ahmadov, R., Tolk, I. 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

Schele, M., Kaminski, T., Knorr, W., Vollbeck, M., Wu, M., Ferrazoli, 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

Schub, 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
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 ringdown laser spectrometers measuring CH4, CO2, and δ13CCH4 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/ORNLDAAC/1568

Sargent, M., Barron, Y., Neuhorn, Y., Huyta, 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/10.1340/2tellusb.v55i2.16731