Constraining Northern Carbon Fluxes with Atmospheric Carbon Dioxide Observations

by

Morgan Loechli

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied Physics)
in the University of Michigan
2023

Doctoral Committee:

Associate Professor Gretchen Keppel-Aleks, Chair
Associate Professor Jeremy N. Bassis
Associate Professor Eric A. Kort
Professor Perry Samson
Professor Allison L. Steiner

Morgan M. Loechli

morgamic@umich.edu

ORCID iD: 0000-0001-5229-6550

Acknowledgements

This work would not have been possible without the guidance and support of Dr. Gretchen Keppel-Aleks and Dr. Britton Stephens. I would also like to acknowledge my friends and family members, especially my husband, Christopher, and my mother, Terry, for continually being my cheerleaders.

Table of Contents

List of Tables

List of Figures

Abstract

Accurate estimates of land carbon fluxes at continental and regional spatial scales require increased understanding of site-to-site differences in seasonal carbon exchange and robust benchmarks against which to evaluate model performance. In this thesis, we use measurements of atmospheric CO_2 to develop a new observationally-based metric related to the land carbon flux against which to evaluate model performance and we probe land-atmosphere carbon exchange across spatial and temporal scales to increase understanding of the spatial distribution of carbon sources and sinks.

First, the northern extratropical growing season net flux (GSNF) is estimated using aircraft profiles of CO_2 measured over the remote oceans. This GSNF is shown to be a robust model benchmark. The northern extratropical GSNF is estimated to be 5.7 ± 0.3 Pg C and coupled model intercomparison project phase 5 (CMIP5) and phase 6 (CMIP6) models are shown to underestimate the GSNF and overestimate the growing season length on average when compared to the observations. This result provides a new and robust observational target of large-scale land carbon flux for prognostic model evaluation.

Second, an emergent constraint approach is applied to prognostic model GSNF to estimate northern extratropical annual fluxes for gross primary productivity (GPP), heterotrophic respiration (RH), and net primary productivity (NPP). Our GSNF-constrained value of 56 ± 15 Pg C for GPP is 8 Pg C larger than a commonly used estimates derived from upscaled flux

towers. Our larger hemispheric GPP estimate in comparison to that from upscaled flux towers indicates that estimates for global GPP may be on the higher end of the current range.

Third, we estimate the northern hemisphere seasonal cycle by upscaling the Total Carbon Column Observing Network (TCCON) to determine how representative the network is of the northern hemisphere. The results indicate that the TCCON is not yet representative of the northern hemisphere. This is likely due to sparse coverage in the high-latitudes. Increasing the coverage and of the network may improve the representativeness of the network.

Lastly, to understand spatial heterogeneity in seasonal cycle amplitudes, we delved into diurnal cycle information from TCCON sites and spatially resolved variations in productivity from space-based data. We saw that saw that nearly all of the variability in the seasonal cycle can be explained by mean March potential temperature at 700 hPa, mean December diurnal cycle amplitude, and maximum annual solar induced chlorophyll fluorescence at a given site.

Chapter 1 : **Introduction**

1.1 Climate Change

Climate change, the long-term increase in global temperatures and associated modification in weather patterns due to the emission of greenhouse gasses (GHGs) to the atmosphere, is one of humanity's most significant challenges. Climate change has already negatively impacted ecosystems, biodiversity, and people beyond natural climate variation (Pörtner et al., 2022). Many lifeforms and ecosystems struggle to adapt to the changing climate leading to an increased rate of extinction (Aitken et al., 2008; Bellard et al., 2012; Halsch et al., 2021; Poloczanska et al., 2013; Raven and Wagner 2021; Reid et al., 2019; Visser 2007; Walsh et al., 2019) and one study by Thomas et al. (2004) predicts 15-37% of species in their sampled regions will go extinct by 2050 under mid-range climate-warming scenarios.

The change in climate causing these extinctions is driven by what is known as the "greenhouse effect". Shortwave radiation from the Sun is transmitted through the atmosphere and absorbed by the Earth. Outgoing longwave radiation emitted by the Earth is partially absorbed by atmospheric GHGs, such as water vapor, carbon dioxide (CO_2), and methane in the Earth's atmosphere. This radiation is then re-emitted, increasing the amount of radiation absorbed at Earth's surface. To achieve energy balance, the surface temperature must increase to increase the total emitted energy consistent with the Stefan-Boltzmann law. Water vapor is Earth's dominate greenhouse gas and is naturally occurring. Other greenhouse gases, such as CO_2, may occur naturally or may be anthropogenic (human caused).

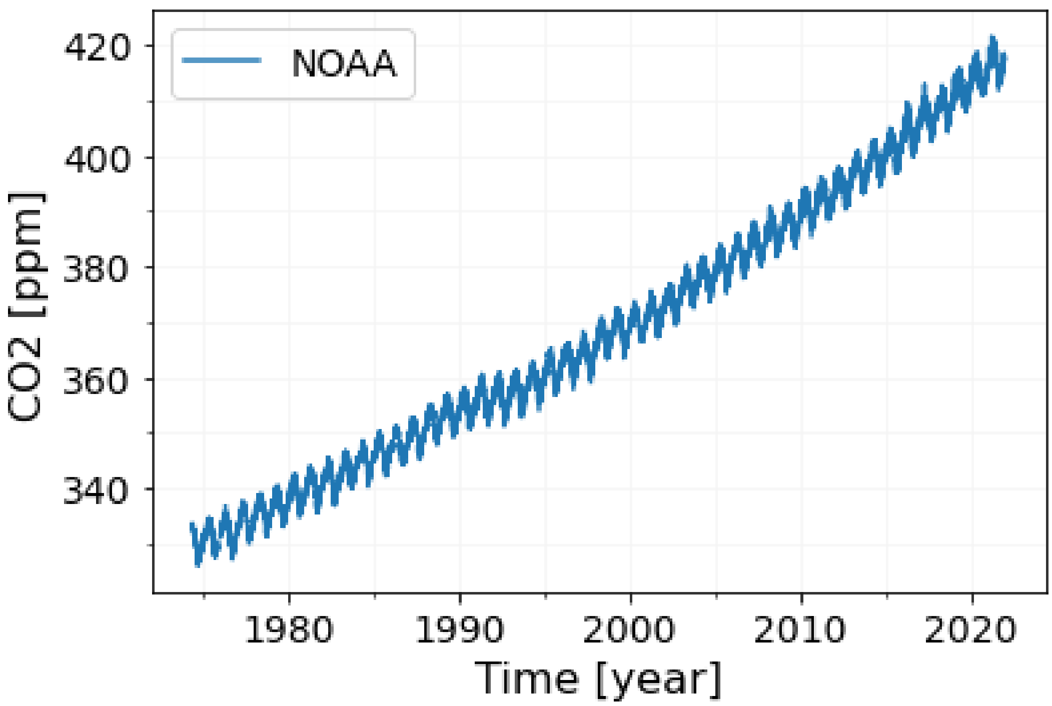


Figure 1.1 Monthly average atmospheric CO_2 concentration at Mauna Loa, Hawaii. The monthly data are from surface measurements from NOAA and the data can be found at https://gml.noaa.gov/ccgg/trends/data.html.(Dr. Pieter Tans, NOAA/GML (gml.noaa.gov/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/).

Without the greenhouse effect, the planet would be too cold to sustain life; however, human activities, such as the burning of fossil fuels, have increased the amount of GHGs in the atmosphere at a rapid rate. The rate of increase is determined not only by fossil fuel emissions, but also by the removal from the atmosphere by the land and oceans. Measurements by the National Oceanic and Atmospheric Administration (NOAA) have shown an increase of atmospheric CO_2 concentration of approximately 100 ppm over the past 60 years (Figure 1.1).

CO_2 is the primary anthropogenic GHG (Bruhwiler et al., 2018; Hansen et al., 1981). The impact a GHG, such as CO_2, has on the climate can be quantified by its radiative forcing. The

Intergovernmental Panel on Climate Change (IPCC) define radiative forcing as the measure of influence a particular factor has on the amount of downward directed radiant energy impacting Earth's surface. The 2013 IPCC assessment estimates a total of approximately 2.3 W/m^2 of radiative forcing due to human activity since pre-industrial times (Ramaswamy et al., 2019). The increase in atmospheric CO_2 between 2000 and 2010 alone was responsible for an increase of 0.2 W/m^2 of radiative forcing (Feldman et al., 2015). This radiative forcing contributes to increasing temperatures and according to the 2022 Global Climate Report from NOAA's National Centers for Environmental Information, the global average temperature in 2022 was 1.55 ° F above the average global temperature for the 20th century.

1.2 Carbon Cycle

Atmospheric CO_2 has a large impact on climate through radiative forcing. The total amount of CO_2 in the atmosphere is determined by emissions as well as uptake by the land and ocean. Thus, understanding when, where, and why carbon enters and leaves the atmosphere is necessary for understanding and simulating climate. On geologic timescales of hundreds of millions to billions of years, the geosphere acts as a reservoir of carbon (Berner 1999, Dasgupta 2013). Carbon enters and exits the geologic reservoir on timescales of millennia unless perturbed by human activity (Pearson and Palmer 2000). On timescales of tens to thousands of years, carbon cycles through the atmosphere, biosphere, and oceans (Berner 1999, Dasgupta 2013). This cycling is governed by processes at the interplay of physics, biology and chemistry. For example, carbon enters the atmosphere at the Earth's surface through respiration of living organisms and exits the atmosphere when dissolved at the ocean's surface.

In the context of anthropogenic climate change, we are primarily concerned with the relatively fast exchange of carbon between the atmospheric, biospheric, and oceanic reservoirs, and we seek to understand the processes that move carbon within and between these reservoirs (Fig 1.2, Ciais et al., 2013). Since the Industrial Revolution, humans have been perturbing the carbon cycle by burning fossil fuels, which rapidly moves carbon from the geologic reservoir into the atmosphere. Around half of the CO_2 released by the combustion of fossil fuel stays in the atmosphere. Friedlingstein et al. (2022) estimate the atmospheric CO_2 growth rate was 47% of total emissions for 2011-2020. The remaining CO_2 is taken up and stored in the biospheric and oceanic sinks, which were estimated in the same study to take on 29% and 26% of total emissions respectively for that decade.

The magnitude of exchange, or what we call strength, of the ocean and land carbon sinks may change as atmospheric CO_2 rises and climate changes. The land is estimated to be a net sink of over 3 PgC yr^{-1} (Fig. 1.2, Friedlingstein et al., 2022). The size of this flux is largely determined by photosynthesis, the primary way the land removes carbon from the atmosphere, and respiration, the primary way the land emits carbon into the atmosphere. It is estimated that approximately 130 PgC enters the atmosphere each year due to respiration and fires (Fig. 1.2, Friedlingstein et al., 2022), and 130 PgC is removed from the atmosphere each year due to photosynthesis (Fig. 1.2, Friedlingstein et al., 2022).

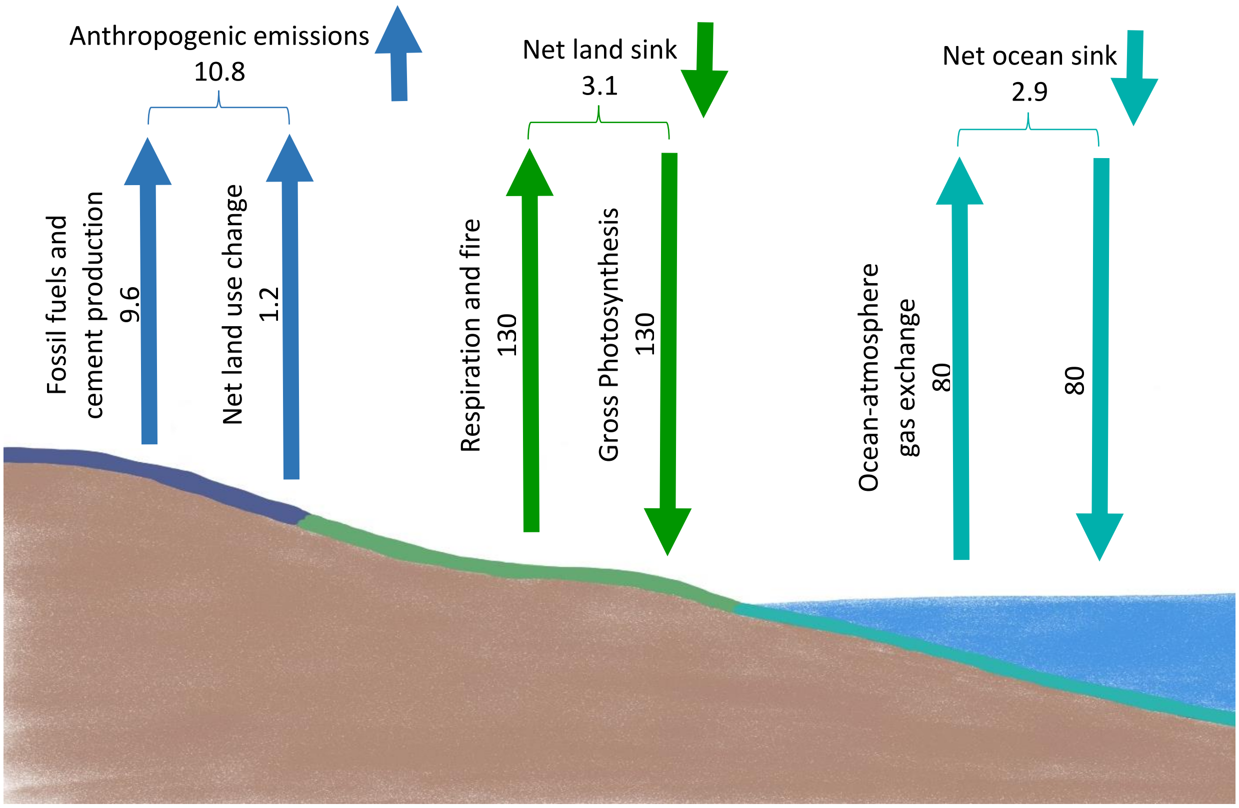


Figure 1.2 Simplified diagram of the global carbon cycle adapted from Figure 2 of Friedlingstein et al., 2022. Numbers represent annual carbon fluxes in PgC yr^{-1} and arrows indicate the direction of the flux.

As emissions have increased over the past century, so too has the strength of the land and ocean sinks, maintaining uptake of approximately half of total emissions. However, previous studies have shown that the current trend of increasing CO_2 uptake by the land and ocean sinks may not continue indefinitely (e.g. Ballantyne et al., 2012; Friedlingstein et al., 2014; Green et al., 2019). Elevated concentrations of CO_2 in the atmosphere can enhance rates of photosynthesis and lead to an increase in land uptake. For example, up to 60% of the present land sink may be due to increasing atmospheric CO_2 (Schimel et al., 2015). However, this “fertilization effect” may be limited in the future by factors such as nutrient availability (Norby et al., 2010), drought (Cooley et al., 2018), and temperature (Yin et al. 2018). The potential change in the land sink under a changing climate highlights the need for an increased mechanistic understanding of the spatial and temporal distribution of CO_2 sources and sinks in order to accurately estimate atmospheric CO_2 and forecast climate.

On the global scale, the land is currently a net sink for CO_2. However, there are large regional variations in carbon fluxes. For example, a regional study by Piao et al. (2012) showed an increase in carbon storage in southern China from 1990-2009 but a decrease in carbon storage in northern East Asia over the same time period.

These fluxes also exhibit seasonal and diurnal variations due to changes in factors such as temperature, light availability, and moisture. For example, the net uptake of CO_2 by terrestrial ecosystems in the Northern Hemisphere is greatest in the summer months when photosynthesis rates are highest. Northern Hemisphere terrestrial ecosystems act as a net source of CO_2 in the winter months when photosynthesis is limited by cold temperatures and shorter days (Baldocchi, 2008; Zhang et al., 2015). Similarly, carbon fluxes will vary over the course of the day. During the daytime, there is ample sunlight available for photosynthesis and plants take up more CO_2 from the atmosphere than they release through respiration. However, at night, photosynthesis ceases while respiration continues and the ecosystem becomes a net source of CO_2 into the atmosphere (e.g. Baldocchi, 2008; Zhang et al., 2015).

These spatial, seasonal, and diurnal variations in land carbon fluxes have important implications for understanding the global carbon cycle and how it will change under a changing climate. A variety of methods are used to measure and model carbon fluxes at different spatial and temporal scales, providing a comprehensive picture of the global carbon budget and allowing for prediction of future climate.

1.3 Climate modeling

Earth system models (ESMs) are valuable tools that can be used to simulate the behavior of the climate system, allowing quantification of natural and human-caused variations to Earth's climate historically and into the future. ESMs represent an advance over climate models in that that fully couple important components of the climate system, including atmosphere, oceans, and terrestrial ecosystems, including by simulating movement of carbon through the Earth system in accordance with feedbacks such as those described above. While many early climate models had the goal of predicting physical climate, current ESMs attempt to predict future climate by accurately simulating all physical, chemical, and biological processes relevant to the carbon cycle. Prognostic climate models must accurately simulate carbon exchange (Bonan & Doney, 2018; Friedlingstein, 2015) to ensure the simulated amount of CO_2 and resulting radiative forcing, are internally consistent and a reasonable expectation for a given scenario (Rogelj et al., 2012; Wigley et al., 1992); however, large differences exist between model projections. For example, the ensemble of Coupled Model Intercomparison Project phase 5 (CMIP5) models shows biases in historical atmospheric CO_2 increase as large as 20%, likely due to differences in the simulation of ocean and land fluxes (Friedlingstein et al., 2014).

This uncertainty can be reduced by assimilating knowledge of previous and current fluxes of carbon dioxide. Observed fluxes provide data for model calibration and targets against which models can be benchmarked. For example, when Friend et al. (2007) used flux measurements made over four vegetation types to calibrate the land-surface scheme of a global climate model, they saw a significant improvement in simulated climate by that model. More recently, Keppel-Aleks et al. (2013) compared atmospheric CO_2 simulations from the Community Earth System Model-Biogeochemisty (CESM1-BGC) to surface, aircraft, and column observations. They found that the model underestimated the atmospheric CO_2 mean annual cycle in the Northern Hemisphere over the historical period, indicating that the net flux in the land component of CESM during the growing season was too weak. One large-scale effort to reduce model uncertainty is the International Land Model Benchmarking (ILAMB) system, which assesses model performance by comparing model output to observational data sets (Collier et al., 2019). The ILAMB system compares multiple models with observations simultaneously which allows relative differences among models and model versions to be quickly identified. Improving models requires fully understanding the spatial and temporal distribution of carbon sources and sinks, as-well-as the factors that influence the strength of those sources and sinks, so that there are good data to use in model calibration and evaluation.

1.4 Measuring and Estimating Carbon Fluxes

Considering that atmospheric CO_2 is dependent on surface fluxes, it is possible to gain information about land-atmosphere carbon exchange from atmospheric CO_2 observations. Measurements of atmospheric CO_2 can be decomposed into three components, a long-term trend, a seasonally-varying component, and a residual. Keeling (1960) found that the long-term trend in atmospheric CO_2, discussed above and shown in figure 1.1, is primarily driven by anthropogenic fossil fuel emissions and that the seasonally-varying component of atmospheric CO_2 is largely driven by terrestrial biospheric carbon fluxes. There is an increase in atmospheric CO_2 during winter months as respiration outpaces photosynthesis. During summer, the net exchange from respiration and photosynthesis is reversed as plants grow, resulting in a decrease in atmospheric

CO_2. As a result of this relationship between atmospheric CO_2 and terrestrial uptake during the growing season, the study of the seasonal cycle can give insight into land surface fluxes. For example, Bolin and Keeling (1963) analyzed the meridional seasonal cycle of atmospheric CO_2 and showed the atmospheric growth rate was smaller than expected from fossil fuel emissions, which they attributed to an ocean carbon sink and which made way for the proposal of a land carbon sink (Tans 1990). A 1985 study by Bacastow et al. concluded that the observed increase in the seasonal amplitude of atmospheric CO_2 at Mauna Loa, Hawaii was most likely due, at least in part, by an increase in plant activity. More recently, a mass-balance analysis incorporating the global mass rate of increase of atmospheric CO_2 and emissions estimates verified that land and ocean sinks have increased in strength over the last 50 years (Ballantyne et al., 2012).

Ideally, we could measure the flux of CO_2 at the atmosphere-ocean interface and at the atmosphere-land interface; however, the vastness and heterogeneity of the Earth's surface make it challenging to get these types of measurements or to robustly interpolate between sites where fluxes are measured (Bastviken et al. 2022). For example, the technique of eddy-covariance, which directly measures net terrestrial ecosystem-atmosphere CO_2 exchange by analyzing high-frequency wind and scalar atmospheric data series (Baldocchi 2003; Goulden et al., 1996), is limited to small spatial scales (Wang et al., 2016). One method of upscaling eddy-covariance flux measurements is presented by Kumar et al. (2016) but the accuracy of the estimates were closely tied with how well sampled an ecoregion was, with regions where sampling was sparse or nonexistent having low accuracy and high uncertainty. Another approach has been to use machine learning to upscale flux measurements made at sites within the FLUXNET network of eddy covariance towers to the globally gridded FLUXCOM product (Jung et al., 2019). While the FLUXCOM product is able to successfully reproduce the carbon cycle at the global scale, the spatial distribution and interannual variability of carbon fluxes found using the FLUXCOM product are not yet reliable (Jung et al., 2020). The difficulty in making reliable observations of carbon fluxes at regional and hemispheric spatial scales motivates the use of models in estimating carbon fluxes.

1.5 Inverse Modeling

The atmospheric concentration of CO_2 at any location is modified both by surface fluxes and atmospheric transport. If atmospheric transport can be well simulated, measurements of atmospheric CO_2 taken in the boundary layer can be used to infer surface fluxes. Previous studies have inferred fluxes on regional and continental scales by observing the concentration distribution of atmospheric CO_2 and assuming information about atmospheric transport (e.g. Bolin and Keeling, 1963; Enting and Mansbridge, 1989; Hyson and Pearman, 1980). Beginning in the 1980s, 3-D CO_2 tracer transport studies were performed (e.g. Fung et al., 1983; Keeling et al., 1989; Tans et al. 1990), with many aiming to evaluate estimated biospheric fluxes by comparing simulated and measured values of atmospheric CO_2. Keeling et al. (1989) found that the model which they evaluated tended to overestimate the seasonal cycle compared to observations and Tans et al. (1990) concluded there must be a northern land carbon sink, spurring a search for said land sink.

These tracer transport studies were the precursor to inversion studies which aimed to estimate land and ocean fluxes. Under an inversion system, models start with some previously assumed information about fossil fuel emissions, ocean flux, and residual (representing the land flux). They then simulate CO_2 concentrations forward in time at the resolution of the atmospheric transport model, typically 1-5 degrees in latitude and longitude. These simulated concentrations are compared to observations and the fluxes are adjusted so that the simulated spatial and temporal variations in CO_2 reasonably match those in the observations (Ciais et al., 2010).

Unfortunately, these optimized fluxes have been shown to be sensitive to the way transport is simulated (e.g. Baker et al., 2006; Gerbig et al., 2008; Gurney et al., 2003; Gurney et al., 2004; Schuh et al., 2019; Verma et al., 2017). The Atmospheric Carbon Cycle Inversion Intercomparison (TransCom 3) project allowed for the comparison of results from an ensemble of models that used the same observations but different atmospheric transport. One result from the TransCom 3 project showed that uptake of CO_2 in large regions, such as the Northern Hemisphere, differed across models due to differences in the simulation of seasonal terrestrial exchange of CO_2 (Gurney et al., 2002). Another result of the TransCom 3 project showed that flux estimates depended on vertical CO_2 gradients, which led to differences in the partitioning of the northern extratropical (NET) and tropical (Tr) land sinks (Stephens et al., 2007). A more recent model Intercomparison project showed that while models are converging on the partitioning of the land exchange, there are still large disagreements on ocean-land flux partitioning (Gaubert et al., 2019). While these models are improving, the differences among transport models remain the largest source of uncertainty in flux estimates made using these models (Schuh et al., 2019).

1.6 Summary and Thesis Overview

Despite the current limitations, we have learned quite a lot over the last 60 years since the first reliable measurements of atmospheric CO_2 were made by Keeling (1960). We now understand that atmospheric CO_2 is increasing due to anthropogenic fossil fuel emissions and that the increase is not as fast as expected due to uptake by land and ocean sinks (Bolin and Keeling, 1963; Bacastow et al., 1985). Land-atmosphere carbon exchange leaves an imprint on atmospheric CO_2 in the form of a seasonally varying component of observed CO_2 (Keeling, 1960). Models have been created that are able to reasonably simulate the Earth's climate (Randall et al., 2007; Flato et al., 2013). Reasonably accurate estimates of carbon fluxes at the local scale can be made using techniques such as eddy-covariance, and at the hemispheric and global scale using atmospheric CO_2 measurements and transport models.

Yet, even with the advancements that have been made, there is still a lot to learn. There is currently difficulty in reliably estimating carbon fluxes at continental, regional, and subregional scales, especially in the tropics (Bruhwiler et al., 2011), and there are still large uncertainties in model simulations of land carbon fluxes (Friedlingstein et al., 2014). One way to address these issues is to use measurements of atmospheric CO_2 to develop new metrics related to the land carbon flux against which to evaluate model performance and to probe land-atmosphere carbon exchange across spatial and temporal scales to increase understanding of the spatial and temporal distribution of carbon sources and sinks.

Integrating CO_2 observations into analyses of flux estimates and model benchmarking is needed to improve climate simulations. In the last 15 years, there has been an increase in measurements of the total abundance of CO_2 in the atmospheric column using both ground-based and space-based instruments. Including measurements of the column-averaged dry-air mole fraction of CO_2 (XCO_2) in inversion studies decreases uncertainty in land flux estimates (e.g. Basu et al. 2018; Byrne et al., 2020). Vertical profiles from aircraft campaigns also provide an opportunity to decrease uncertainty in flux estimates when assimilated into inversion systems (e.g. Verma et al., 2017). Here, we propose that these observations of atmospheric CO_2 can be used to increase understanding of seasonal land-atmosphere carbon exchange and to create meaningful metrics against which to evaluate model performance without relying heavily on the use of inverse models, which introduce uncertainty. The aims of this work are as follows:

- We use CO_2 profiles measured over the remote oceans to infer the Northern extratropical growing season net flux (GSNF) and show that this flux can be used to benchmark model performance.

- We employ an emergent constraint approach to estimate Northern extratropical annual fluxes for gross primary production (GPP), due to photosynthesis, heterotrophic respiration (RH), due to decomposition of organic matter by microorganisms, and net primary production (NPP), which is GPP minus plant respiration.

- We investigate the representativeness of a network of ground-based instruments measuring column-averaged dry air mole fractions of CO_2 (XCO_2).

- We utilize ground-based and space-based measurements of XCO_2 to connect across temporal and spatial scales to explain differences in site-to-site seasonal cycles and gain understanding of sub-regional carbon fluxes.

1.7 References

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, *1*(1), 95–111. https://doi.org/10.1111/j.1752-4571.2007.00013.x

Bacastow, R. B., Keeling, C. D., & Whorf, T. P. (1985). Seasonal amplitude increase in atmospheric CO2 concentration at Mauna Loa, Hawaii, 1959 - 1982. *Journal of Geophysical Research*, *90*(D6), 10529–10540. https://doi.org/10.1029/JD090iD06p10529

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), n/a-n/a. https://doi.org/10.1029/2004GB002439

Baldocchi, D. D. (2003). Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology*, *9*(4), 479–492. https://doi.org/10.1046/j.1365-2486.2003.00629.x

Baldocchi, D. (2008). “Breathing” of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Australian Journal of Botany*, *56*(1), 1. https://doi.org/10.1071/BT07151

Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., & White, J. W. C. (2012). Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature*. https://doi.org/10.1038/nature11299

Bastviken, D., Wilk, J., Duc, N. T., Gålfalk, M., Karlson, M., Neset, T.-S., Opach, T., et al. (2022). Critical method needs in measuring greenhouse gas fluxes. *Environmental Research Letters*, *17*(10), 104009. https://doi.org/10.1088/1748-9326/ac8fa9

Basu, S., Baker, D. F., Chevallier, F., Patra, P. K., Liu, J., & Miller, J. B. (2018). The impact of transport model differences on CO2 surface flux estimates from OCO-2 retrievals of column average CO2. *Atmospheric Chemistry and Physics*, *18*(10). https://doi.org/10.5194/acp-18-7189-2018

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. *Ecology Letters*, *15*(4), 365–377. https://doi.org/10.1111/j.1461-0248.2011.01736.x

Berner, Robert A. "A new look at the long-term carbon cycle." *Gsa Today* 9.11 (1999): 1-6.

Bolin, B., & Keeling, C. D. (1963). Large-scale atmospheric mixing as deduced from the seasonal and meridional variations of carbon dioxide. *Journal of Geophysical Research*, *68*(13), 3899–3920. https://doi.org/10.1029/JZ068i013p03899

Bonan, G. B., & Doney, S. C. (2018). Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models. In *Science*. https://doi.org/10.1126/science.aam8328

Bruhwiler, L. M. P., Michalak, A. M., and Tans, P. P.: Spatial and temporal resolution of carbon flux estimates for 1983–2002, Biogeosciences, 8, 1309–1331, https://doi.org/10.5194/bg-8-1309-2011

Bruhwiler, L., Michalak, A. M., Birdsey, R., Huntzinger, D. N., Fisher, J. B., & Miller, J. (2018). *Chapter 1: Overview of the Global Carbon Cycle. Second State of the Carbon Cycle Report*. https://doi.org/10.7930/SOCCR2.2018.Ch1

Byrne, B., Liu, J., Lee, M., Baker, I., Bowman, K. W., Deutscher, N. M., Feist, D. G., 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). https://doi.org/10.1029/2019JD032029

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., Cotrim da Cunha, L., Cox, P. M., et al. Global Carbon and other Biogeochemical Cycles and Feedbacks, in: Climate Change 2021: The Physical Science Basis, Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 673–816, https://doi.org/10.1017/9781009157896.007, 2021.

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., & Ramonet, M. (2010). Atmospheric inversions for estimating CO2 fluxes: Methods and perspectives. *Climatic Change*. https://doi.org/10.1007/s10584-010-9909-3

Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Quéré, R.B. Myneni, S. Piao and P. Thornton, 2013: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical

Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., et al. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *Journal of Advances in Modeling Earth Systems*. https://doi.org/10.1029/2018MS001354

Cooley, S. R., Moore, D. J. P., Alin, S. R., Butman, D., Clow, D. W., French, N. H. F., et al. (2018). *Chapter 17: Biogeochemical Effects of Rising Atmospheric Carbon Dioxide. Second State of the Carbon Cycle Report*. 690–727.https://doi.org/10.7930/SOCCR2.2018.Ch17

Dasgupta, R. (2013). Ingassing, Storage, and Outgassing of Terrestrial Carbon through Geologic Time. *Reviews in Mineralogy and Geochemistry*, *75*(1), 183–229. https://doi.org/10.2138/rmg.2013.75.7

Dlugokencky, E. and Tans, P.: Trends in atmospheric carbon dioxide, National Oceanic and Atmospheric Administration, Global Monitoring Laboratory (NOAA/GML), http://www.gml.noaa.gov/gmd/ccgg/trends/global.html, last access: 25 September 2022.

Enting, I. G., & Mansbridge, J. v. (1989). Seasonal sources and sinks of atmospheric CO $_2$ Direct inversion of filtered data. *Tellus B*, *41B*(2), 111–126. https://doi.org/10.1111/j.1600-0889.1989.tb00129.x

Feldman, D. R., Collins, W. D., Gero, P. J., Torn, M. S., Mlawer, E. J., & Shippert, T. R. (2015). Observational determination of surface radiative forcing by CO2 from 2000 to 2010. *Nature*, *519*(7543), 339–343. https://doi.org/10.1038/nature14240

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, et al. 2013: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

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

Friedlingstein, P. (2015). Carbon cycle feedbacks and future climate change. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. https://doi.org/10.1098/rsta.2014.0421

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., et al..: Global Carbon Budget 2021, Earth Syst. Sci. Data, 14, 1917–2005, https://doi.org/10.5194/essd-14-1917-2022, 2022.

Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogée, J., Rödenbeck, C., et al. (2007). FLUXNET and modelling the global carbon cycle. *Global Change Biology*. https://doi.org/10.1111/j.1365-2486.2006.01223.x

Fung, I., Prentice, K., & Matthews, E. (1983). Three-dimensional tracer model study of atmospheric CO2: Response to seasonal exchanges with the terrestrial biosphere. *Journal of Geophysical Research*, *88*(C2), 1281–1294. https://doi.org/10.1029/JC088iC02p01281

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., Patra, P. K., et al. (2019). Global atmospheric CO2 inverse models converging on neutral tropical land exchange but diverging on fossil fuel and atmospheric growth rate. *Biogeosciences Discussions*, 1–25. https://doi.org/10.5194/bg-2018-384

Gerbig, C., Körner, S., & Lin, J. C. (2008). Vertical mixing in atmospheric tracer transport models: error characterization and propagation. *Atmospheric Chemistry and Physics*, *8*(3), 591–602. https://doi.org/10.5194/acp-8-591-2008

Goulden, M. L., Munger, J. W., Fan, S.-M., Daube, B. C., & Wofsy, S. C. (1996). Measurements of carbon sequestration by long-term eddy covariance: methods and a critical evaluation of accuracy. *Global Change Biology*, *2*(3), 169–182. https://doi.org/10.1111/j.1365-2486.1996.tb00070.x

Green, J. K., Seneviratne, S. I., Berg, A. M., Findell, K. L., Hagemann, S., et al. (2019). Large influence of soil moisture on long-term terrestrial carbon uptake. In *Nature* (Vol. 565, Issue 7740, pp. 476–479). Nature Publishing Group. https://doi.org/10.1038/s41586-018-0848-x

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*. https://doi.org/10.1038/415626a

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., et al. (2003). TransCom 3 CO $_2$ inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. *Tellus B: Chemical and Physical Meteorology*, *55*(2), 555–579. https://doi.org/10.3402/tellusb.v55i2.16728

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., et al. (2004). Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochemical Cycles*, *18*(1), n/a-n/a. https://doi.org/10.1029/2003GB002111

Halsch, C. A., Shapiro, A. M., Fordyce, J. A., Nice, C. C., Thorne, J. H., Waetjen, D. P., et al. (2021). Insects and recent climate change. *Proceedings of the National Academy of Sciences*, *118*(2). https://doi.org/10.1073/pnas.2002543117

Hansen, J., Johnson, D., Lacis, A., Lebedeff, S., Lee, P., Rind, D., & Russell, G. (1981). Climate impact of increasing atmospheric carbon dioxide. *Science*. https://doi.org/10.1126/science.213.4511.957

Hyson, P., Fraser, P. J., & Pearman, G. I. (1980). A two-dimensional transport simulation model for trace atmospheric constituents. *Journal of Geophysical Research*, *85*(C8), 4443. https://doi.org/10.1029/JC085iC08p04443

Jung, M., Koirala, S., Weber, U. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci Data* **6**, 74 (2019). https://doi.org/10.1038/s41597-019-0076-8

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020

Keeling, C. D. (1960) The Concentration and Isotopic Abundances of Carbon Dioxide in the Atmosphere, Tellus, 12:2, 200-203, DOI: 10.3402/tellusa.v12i2.9366

Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R., et al.: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, Tellus A., 28, 538–551, https://doi.org/10.1111/j.2153-3490.1976.tb00701.x, 1976.

Keeling, C.D., Piper, S.C. and Heimann, M. (1989). A three-dimensional model of atmospheric CO2 transport based on observed winds: 4. Mean annual gradients and interannual variations. In *Aspects of Climate Variability in the Pacific and the Western Americas*, D.H. Peterson (Ed.). https://doi.org/10.1029/GM055p0305

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., et al. (2013). Atmospheric carbon dioxide variability in the community earth system model: Evaluation and transient dynamics during the twentieth and twenty-first centuries. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-12-00589.1

Kumar, J., Hoffman, F., Hargrove, W., & Collier, N. (2016). Understanding the representativeness of FLUXNET for upscaling carbon flux from eddy covariance measurements. *Earth System Science Data Discussions*, 1–25. https://doi.org/10.5194/essd-2016-36

Masarie, K. A. and Tans, P. P.: Extension and integration of atmospheric carbon dioxide data into a globally consistent measurement record, J. Geophys. Res., 100, 11593, https://doi.org/10.1029/95JD00859, 1995.

McLaughlin, J. F., Hellmann, J. J., Boggs, C. L., & Ehrlich, P. R. (2002). Climate change hastens population extinctions. *Proceedings of the National Academy of Sciences*, *99*(9), 6070–6074. https://doi.org/10.1073/pnas.052131199

NOAA National Centers for Environmental Information (2023). State of the Climate: Global Climate Report for 2022. Accessed February 10, 2023, from https://www.ncei.noaa.gov/access/monitoring/monthly-report/global/202213.

Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E., & McMurtrie, R. E. (2010). CO_2 enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences*, *107*(45), 19368–19373. https://doi.org/10.1073/pnas.1006463107

Pearson, P. N., & Palmer, M. R. (2000). Atmospheric carbon dioxide concentrations over the past 60 million years. *Nature*, *406*(6797), 695–699. https://doi.org/10.1038/35021000

Piao, S. L., Ito, A., Li, S. G., Huang, Y., Ciais, P., Wang, X. H., Peng, S. S., Nan, H. J., et al. The carbon budget of terrestrial ecosystems in East Asia over the last two decades, Biogeosciences, 9, 3571–3586, https://doi.org/10.5194/bg-9-3571-2012, 2012.

Poloczanska, E. S., Brown, C. J., Sydeman, W. J., Kiessling, W., Schoeman, D. S., et al. (2013). Global imprint of climate change on marine life. *Nature Climate Change*, *3*(10), 919–925. https://doi.org/10.1038/nclimate1958

Pörtner, H.-O., D.C. Roberts, H. Adams, I. Adelekan, C. Adler, R. Adrian, et al. 2022: Technical Summary. [H.-O. Pörtner, D.C. Roberts, E.S. Poloczanska, K. Mintenbeck, M. Tignor, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem (eds.)]. In: *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp. 37–118, doi:10.1017/9781009325844.002.

Price, J. T. and Warren, R.: Literature Review of the Potential of “Blue Carbon” Activities to Reduce Emissions, https://avoid-net-uk.cc.ic.ac.uk/wp-content/uploads/delightful-downloads/2016/03/Literature-review-of-the-potential-of-blue-carbon-activities-to-reduce-emissions-AVOID2-WPE2.pdf (last access: 25 September 2022), 2016.

Ramaswamy, V., Collins, W., Haywood, J., Lean, J., Mahowald, N., Myhre, G., Naik, V., et al. (2019). Radiative Forcing of Climate: The Historical Evolution of the Radiative Forcing

Concept, the Forcing Agents and their Quantification, and Applications. *Meteorological Monographs*, *59*, 14.1-14.101. https://doi.org/10.1175/AMSMONOGRAPHS-D-19-0001.1

Randall, D.A., R.A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, et al. 2007: Cilmate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Raven, P. H., & Wagner, D. L. (2021). Agricultural intensification and climate change are rapidly decreasing insect biodiversity. *Proceedings of the National Academy of Sciences*, *118*(2). https://doi.org/10.1073/pnas.2002548117

Reid, A. J., Carlson, A. K., Creed, I. F., Eliason, E. J., Gell, P. A., Johnson, P. T. J., et al. (2019). Emerging threats and persistent conservation challenges for freshwater biodiversity. *Biological Reviews*, *94*(3), 849–873. https://doi.org/10.1111/brv.12480

Rogelj, J., Meinshausen, M., & Knutti, R. (2012). Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Climate Change*, *2*(4), 248–253. https://doi.org/10.1038/nclimate1385

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1407302112

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., Chevallier, F., et al. (2019). Quantifying the Impact of Atmospheric Transport Uncertainty on CO 2 Surface Flux Estimates. *Global Biogeochemical Cycles*, *33*(4), 484–500. https://doi.org/10.1029/2018GB006086

Schneider, T., O'Gorman, P. A., & Levine, X. J. (2010). WATER VAPOR AND THE DYNAMICS OF CLIMATE CHANGES. *Reviews of Geophysics*, *48*(3), RG3001. https://doi.org/10.1029/2009RG000302

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., et al. (2007). Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO2. *Science*. https://doi.org/10.1126/science.1137004

Tans, P. P., Fung, I. Y., & Takahashi, T. (1990). Observational constraints on the global atmospheric CO2 budget. *Science*. https://doi.org/10.1126/science.247.4949.1431

Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., et al. (2004). Extinction risk from climate change. *Nature*, *427*(6970), 145–148. https://doi.org/10.1038/nature02121

Verma, S., Marshall, J., Gerbig, C., Rödenbeck, C., & Uwe Totsche, K. (2017). The constraint of CO2 measurements made onboard passenger aircraft on surface-atmosphere fluxes: The impact of transport model errors in vertical mixing. *Atmospheric Chemistry and Physics*, *17*(9), 5665–5675. https://doi.org/10.5194/acp-17-5665-2017

Visser, M. E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1635), 649–659. https://doi.org/10.1098/rspb.2007.0997

Walsh, B. S., Parratt, S. R., Hoffmann, A. A., Atkinson, D., Snook, R. R., Bretman, A., & Price, T. A. R. (2019). The Impact of Climate Change on Fertility. *Trends in Ecology & Evolution*, *34*(3), 249–259. https://doi.org/10.1016/j.tree.2018.12.002

Wang, H., Jia, G., Zhang, A., & Miao, C. (2016). Assessment of Spatial Representativeness of Eddy Covariance Flux Data from Flux Tower to Regional Grid. *Remote Sensing*, *8*(9), 742. https://doi.org/10.3390/rs8090742

Wigley, T. M. L., & Raper, S. C. B. (1992). Implications for climate and sea level of revised IPCC emissions scenarios. *Nature*, *357*(6376), 293–300. https://doi.org/10.1038/357293a0

Yin, Y., Ciais, P., Chevallier, F., Li, W., Bastos, A., Piao, S., Wang, T., & Liu, H. (2018). Changes in the Response of the Northern Hemisphere Carbon Uptake to Temperature Over the Last Three Decades. *Geophysical Research Letters*, *45*(9), 4371–4380. https://doi.org/10.1029/2018GL077316

Zhang, L., Sun, R., Xu, Z., Qiao, C., & Jiang, G. (2015). Diurnal and Seasonal Variations in Carbon Dioxide Exchange in Ecosystems in the Zhangye Oasis Area, Northwest China. *PLOS ONE*, *10*(3), e0120660. https://doi.org/10.1371/journal.pone.0120660

Chapter 2 : Evaluating Northern Hemisphere Growing Season Net Carbon Flux in Climate Models Using Aircraft Observations

This chapter has been previously published at Global Biogeochemical Cycles.

Loechli, Morgan & Stephens, Britton & Commane, Roisin & Chevallier, Frederic & McKain, Kathryn & Ralph, Keeling & Morgan, Eric & Patra, P. & Sargent, M. & Sweeney, Colm & Keppel-Aleks, Gretchen. (2023). Evaluating Northern Hemisphere Growing Season Net Carbon Flux in Climate Models Using Aircraft Observations. https://doi.org/10.1029/2022GB007520

Key Points:

- Aircraft observations of atmospheric carbon dioxide concentrations are used to infer the northern extratropical growing season net flux.
- The observations suggest a larger net flux and shorter growing season than simulated in Earth system models.
- An emergent constraint approach is used to estimate productivity and respiration fluxes.

Abstract

Understanding terrestrial ecosystems and their response to anthropogenic climate change requires quantification of land-atmosphere carbon exchange. However, top-down and bottom-up estimates of large-scale land-atmosphere fluxes, including the northern extratropical growing season net flux (GSNF), show significant discrepancies. We develop a data-driven metric for the GSNF using atmospheric carbon dioxide concentration observations collected during the High-Performance Instrumented Airborne Platform for Environmental Research (HIAPER) Pole-to-Pole Observations (HIPPO) and Atmospheric Tomography Mission (ATom) flight campaigns. This aircraft-derived metric is bias-corrected using three independent atmospheric inversion systems. We estimate the northern extratropical GSNF to be 5.7 ± 0.3 Pg C and use it to evaluate net biosphere productivity from the Coupled Model Intercomparison Project phase 5 and 6 (CMIP5 and CMIP6) models. While the model-to-model spread in the GSNF has decreased in CMIP6 models relative to that of the CMIP5 models, there is still disagreement on the magnitude and timing of seasonal carbon uptake with most models underestimating the GSNF and overestimating the length of the growing season relative to the observations. We also use an emergent constraint approach to estimate annual northern extratropical gross primary productivity to be 56 ± 17 Pg C, heterotrophic respiration to be 25 ± 13 Pg C, and net primary productivity to be 28 ± 12 Pg C. The flux inferred from these aircraft observations provides an additional constraint on large-scale, gross fluxes in prognostic Earth system models that may ultimately improve our ability to accurately predict carbon-climate feedbacks.

Plain Language Summary

The exchange of carbon between the land and atmosphere is an important part of the Earth's climate, and this exchange might change due to human-caused climate change. However, estimates of land-atmosphere carbon fluxes made using different techniques do not agree with each other. We use atmospheric carbon dioxide observations collected during two flight campaigns to show that 5.7 Pg C are exchanged between the atmosphere and the land in the northern hemisphere during the summer growing season. This estimate is used to evaluate the performance of two generations of climate prediction models. The newer generation of models show less spread than the older generation, but there is still significant disagreement on the magnitude and timing of land-atmosphere carbon exchange among models. Most models underestimate the growing season net flux and overestimate the length of the growing season. We also use our observational estimate to reduce the spread on component fluxes of carbon exchange, namely uptake by photosynthesis and release by respiration.

2.1 Introduction

Approximately half of the carbon dioxide (CO_2) released annually by the combustion of fossil fuels stays in the atmosphere (Keeling et al., 1976; Schimel et al., 2001; Friedlingstein et al., 2021). The remaining CO_2 is taken up by the terrestrial biosphere and ocean in roughly equal proportion (Khatiwala et al., 2009, Keeling et al., 2014; Sabine et al., 2004). The efficiency of the ocean and land sinks varies with both climate and atmospheric CO_2, representing an important feedback in the climate system (e.g. Ballantyne et al. 2012; Fung et al., 2005; Fernandez-Martinez et al., 2019). The strength of the land sink may be related to the amplitude of the seasonal cycle of atmospheric CO_2 (e.g. Keeling et al., 1996; Randerson et al., 1997) via annual and seasonal imbalances between photosynthesis and respiration. However, the magnitude, and spatial and temporal distributions of gross primary productivity (GPP) and net primary productivity (NPP) vary noticeably among Earth system models (ESMs) (e.g. Hu et al., 2022). Furthermore, models typically underestimate the change in amplitude of seasonal CO_2 exchange in northern land ecosystems over time (e.g. Graven et al., 2013) or underestimate CO_2 uptake in the Northern Hemisphere mid-high latitudes. (e.g. Canadell et al., 2021, fig 5.24).

Multi-model ensembles of coupled carbon-climate models show large differences in their land sink projections, especially for terrestrial carbon uptake (e.g. Arora et al., 2020; Cadule et al., 2010). For example, Friedlinstein et al. (2014) showed that the Coupled Model Intercomparison Project phase 5 (CMIP5) models range between -173 and 758 Pg C in simulations of cumulative land carbon uptake for 1850 to 2100 when forced by RCP8.5. This uncertainty exists in historical simulations where models both overestimate and underestimate the historical atmospheric CO_2 increase by over 20%. These differences are mainly due to uncertainties in the land carbon cycle response, with differences in their cumulative land flux estimates of 214 Pg C for 1850-2005, more than double the differences in their cumulative ocean flux estimates (Friedlingstein et al., 2014).

Quantifying the exchange of carbon between the atmosphere and the land surface at hemispheric and global scales is challenging because the heterogeneity of Earth’s surface makes it difficult to upscale local flux measurements (e.g. Friend et al., 2007; Kumar et al., 2016). Atmospheric inversion, wherein carbon fluxes are estimated from atmospheric CO_2 observations using atmospheric tracer transport models, provides a method to infer large-scale carbon fluxes (e.g. Tans et al., 1990; Ciais et al., 2010; Thompson et al., 2016); however, this method has been shown to be sensitive to uncertainty due to the simulation of vertical transport (Schuh et al., 2019; Stephens et al., 2007; Verma et al., 2017). Atmospheric inversions that rely only on surface observations must accurately represent vertical mixing to estimate CO_2 concentrations aloft. Uncertainty in atmospheric inversion flux estimates can be characterized through the use of observations of the vertical profile of atmospheric CO_2 (e.g. Peiro et al., 2022; Stephens et al., 2007).

Global-scale aircraft observations, such as those made during the High-Performance Instrumented Airborne Platform for Environmental Research (HIAPER) Pole-to-Pole Observations project (HIPPO, 2009-11) and the Atmospheric Tomography Mission (ATom, 2016-18), are representative of large regions and capture the vertical profile of atmospheric CO_2 (Wofsy et al., 2011; Thompson et al, 2021). These campaigns measured the vertical structure of CO_2 in the atmosphere across a range of latitudes and over the full seasonal cycle, and allow for analysis of seasonal changes in hemispheric-scale atmospheric CO_2 (e.g. Jin et al., 2021), which are dominated by land exchange. We use the seasonal cycle of atmospheric CO_2 concentrations measured during the HIPPO and ATom flight campaigns to develop a metric for evaluating the simulation of terrestrial CO_2 exchange in prognostic ESMs.

We derive estimates of the northern hemisphere net land flux integrated over the growing season, or growing season net flux (GSNF), as a benchmark for model evaluation (e.g. Collier et al., 2018). The creation of flux benchmarks allows for a direct comparison of observations and model simulations at the flux level rather than at the concentration level (e.g. Keppel-Aleks et al., 2013), which requires either using an atmospheric transport model or emulator (Liptak et al., 2017) to translate fluxes into atmospheric mole fraction variations. This research explores an alternative approach to formal inverse modeling to constrain net land-atmosphere carbon fluxes at hemispheric scale. We use CO_2 measurements from the HIPPO and ATom flight campaigns to infer the GSNF with only minimal reliance on atmospheric transport models. Thus, our estimated flux is less sensitive to errors in transport simulation and gives more robust insight into prognostic model inconsistencies.

We describe the data sets and methods used to derive GSNF in Section 2.2. We discuss the GSNF and compare to ESM estimates of net biosphere productivity (NBP), GPP, heterotrophic respiration (RH), and net primary productivity (NPP), using output from the Coupled Model Intercomparison Project phase 5 and 6 (CMIP5 and CMIP6) in Section 2.3. This is followed by a discussion of those results in Section 2.4 and conclusions in Section 2.5.

2.2 Data and Methods

2.2.1 Aircraft Observations

We used dry air mole fractions of CO_2 measured in the free troposphere during the HIPPO and ATom aircraft campaigns. HIPPO (Wofsy, 2011; Wofsy et al. 2017) used the NSF/NCAR HIAPER Gulfstream V aircraft to make measurements primarily over the remote Pacific from 87° N to 67° S (Fig. 2.1a) during five campaigns that spanned all four seasons between 2009 and 2011 (Table 2.1). The aircraft flew vertical profiles from near the surface to an altitude of 14 km; typically, a full profile was completed over ~2.2° of latitude (Fig. 2.1b). During these flights, measurements were made of greenhouse gasses and related tracers. CO_2 mole fractions were measured using three different in situ instruments and two whole air samplers: the Harvard Quantum Cascade Laser System (QCLS, Santoni et al., 2014), the Harvard Observations of the Middle Stratosphere (OMS, Daube et al., 2002) instrument, the National Center for Atmospheric Research (NCAR) Airborne Oxygen Instrument (AO2, Stephens et al., 2021), the National Oceanic and Atmospheric Administration (NOAA) Portable Flask Packages (PFP, Sweeney et al., 2015), and the NCAR/Scripps Medusa Whole Air Sampler (Stephens et al., 2021). For our analysis, we used the recommended CO2.X variable, which is derived primarily from QCLS measurements with calibration periods gap-filled using OMS measurements, reported as part per million dry air mole fraction (Wofsy et al., 2017). We used the 10-second merge data product and all CO_2 measurements are reported to be within 0.2 ppm with respect to the WMO X2007 scale (Santoni et al., 2014). The mean bias between QCLS and NOAA flask measurements across all five HIPPO campaigns is 0.11 ppm (Santoni et al., 2014). We used comparisons to the other 4 systems as a measure of analytical uncertainty. We also used observations of N_2O made by QCLS to identify stratospheric samples.

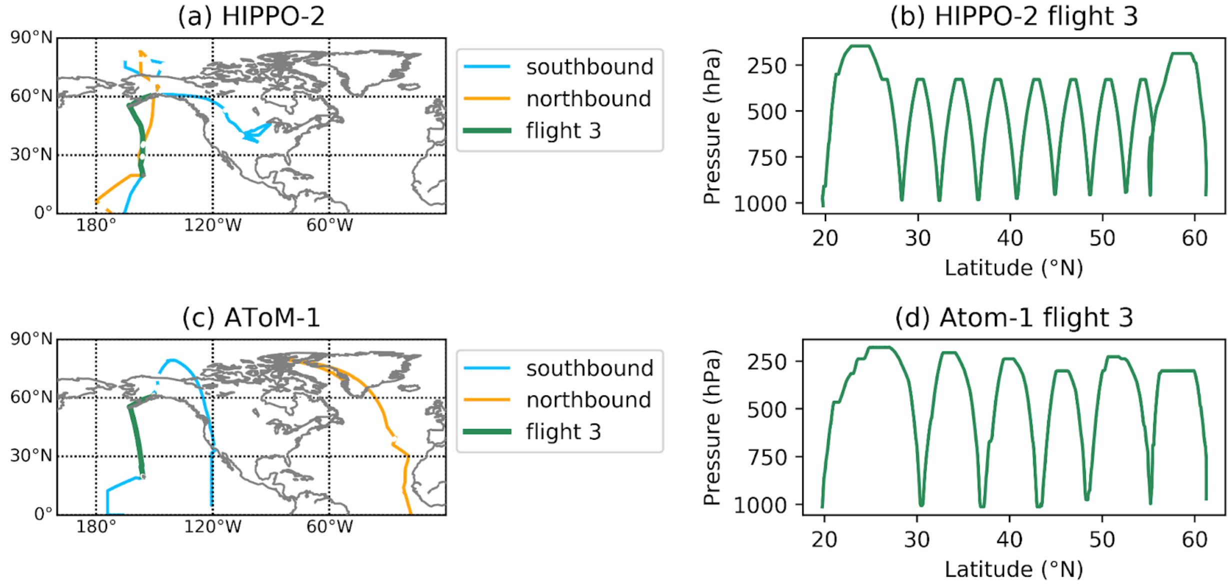


Figure 2.1 Flight paths for (a) HIPPO-2, which flew over the remote Pacific in November 2009, and (c) ATom-1, which flew over the Pacific and Atlantic in August 2016. All other campaigns followed similar flight paths. Flight path with continuous vertical profiling for flight 3, which flew from Anchorage, AK to Kona, HI for (b) HIPPO-2 and (d) AToM-1. All other flights flew a similar path.

ATom (Wofsy et al., 2021; Thompson et al., 2022) is a more recent series of flight campaigns that used the NASA DC-8 aircraft to measure atmospheric trace gas concentrations by traveling south over the Pacific and north over the Atlantic (Fig. 2.1c) and which included a much larger scientific payload. As with HIPPO, a full annual cycle was measured, with flights that occurred in each of the four seasons over a three-year period from 2016 to 2018 (Table 2.1). Flights spanned 83°N to 86°S and sampled vertical profiles from 0.2 to 12 km in altitude (Fig. 2.1d). ATom measured CO_2 using the QCLS, AO2, Medusa, and PFPs similarly to HIPPO but also included a NOAA Picarro instrument. For our analysis, we used the CO2.X variable, which consists of NOAA Picarro measurements gap-filled using QCLS measurements. However during the first two flights of ATom-1, the NOAA Picarro measurements were not reported due to an inlet problem. Similar to HIPPO, we used the 10-second merge data product (Wofsy et al., 2021). To identify and remove stratospheric samples, we used observations of N_2O from QCLS and the NOAA PAN and Trace Hydrohalocarbon ExpeRiment (PANTHER, ATom-1 only). While the WMO CO_2 scale has been recently updated, both the HIPPO and ATom observations used here have been calibrated with respect to the previous, WMO X2007, scale.

Table 2.1 Aircraft data used in this study

Deployment	Northern Hemisphere (Southbound) Dates and Ocean Basin	Northern Hemisphere (Northbound) Dates and Ocean Basin
HIPPO-1	1/8/09-1/16/09, Pacific	1/28/09-1/30/09, Pacific
HIPPO-2	10/31/09-11/7/09, Pacific	11/16/09-11/22/09, Pacific
HIPPO-3	3/24/10-3/31/10, Pacific	4/10/10-4/16/10, Pacific
HIPPO-4	6/14/11-6/22/11, Pacific	7/4/11-7/11/11, Pacific
HIPPO-5	8/9/11-8/24/11, Pacific	9/3/11-9/8/11, Pacific
ATom-1	7/29/16-8/6/16, Pacific	8/17/16-8/23/16, Atlantic
ATom-2	1/26/17-2/3/17, Pacific	2/15/16-2/21/17, Atlantic
ATom-3	9/28/17-10/6/17, Pacific	10/19/17-10/28/17, Atlantic
ATom-4	4/24/18-5/1/18, Pacific	5/14/18-5/21/18, Atlantic

2.2.2 Curtain Averages from Atmospheric Concentrations

The CO_2 observations from all flight campaigns are combined to estimate the average northern extratropical tropospheric CO_2 seasonal cycle (Bent, 2014). We then use a set of transport models to convert the time derivative of this cycle into estimates of northern extratropical terrestrial CO_2 flux. We refer to this process as “bias correction.”

To isolate tropospheric CO_2 signals, we define an upper cutoff of 300 hPa and remove any remaining observations with detectable stratospheric influence using the measured concentration of nitrous oxide (N_2O) and a cutoff value of 319 parts per billion (ppb) after detrending the data to 2009; samples whose N_2O concentration falls below this threshold are removed from the observations (Bent, 2014). We also manually remove outlying samples primarily obtained during takeoffs and landings, to avoid strong local influences from biospheric exchange or fossil emissions. The flights and times filtered are identified in Data Sets S1 and S2. We filter output at the same locations and times for the transport model CO_2 mole fractions simulated along the flight tracks, discussed in section 2.2.3. This stratospheric and local influence filtering removes 2.8% of the observations within the defined domain from the HIPPO and ATom datasets. We do not use observations from the northbound leg of HIPPO-1 because it only extended to 40 ° N, and both QCLS and OMS have been filtered for altitude-dependent biases on these flights.

We then detrend the filtered data by removing the long-term trend in the NOAA Mauna Loa in situ CO_2 mole fraction record (Thoning et al., 2022), found by Seasonal-Trend decomposition using locally estimated scatterplot smoothing (STL, Cleveland et al., 1990) with a 2-year smoothing window. By detrending, the overall annual mean level of flux is removed, leaving only the (relative) seasonal cycle.

We calculate the extratropical mean drawdown by first aggregating the detrended data in latitude and pressure bins. We discretize the atmosphere into bins of 5° in latitude and 50 hPa in pressure, for the latitude range 20°N to 90°N and the pressure range 300 hPa to 1000 hPa. Observations at latitudes south of 20°N are excluded because of the differences in the phasing of the tropical seasonal cycle to that north of 20°N, and observations at pressures below 300 hPa were excluded because measurements were sparse and frequently in the stratosphere. Within each bin, we average all data collected for a given day of the year and then fit a second-order harmonic as a function of day of the year with an offset due to the difference in the annual mean relative to Mauna Loa (Fig. S1). We then generate seasonal time series at daily resolution from the harmonic fits and take the pressure-weighted average of these values for each latitude bin. These partial columns are then integrated over latitude from 20°N to 90°N (Eq. 1), using cosine(lat) weighting to reflect the influence on the zonal volume below 300 hPa at latitudes where observations were made. We call the result of this integration the "curtain average" concentration of atmospheric CO_2 (Bent, 2014).

$$CurtainAvg = \frac{\int_{20°N}^{90°N} \int_{300\,hPa}^{1000\,hPa} aveCO_2(\varphi, P)\, dp \cos\varphi d\varphi}{\int_{20°N}^{90°N} \int_{300\,hPa}^{1000\,hPa} dp \cos\varphi d\varphi} \qquad (1)$$

The curtain average is shown in black in Figure 2.2, and is compared to northbound and southbound legs of each HIPPO and ATom mission where each point is found by filtering, detrending, interpolating and extrapolating to get a full altitude and latitude slice, then taking a pressure and cosine of latitude weighted average (Akima, 1978).

The derivative, found as a finite-difference, of the curtain average concentration fit line with respect to time, then gives the rate of change of CO_2 of this atmospheric volume as a function of the day of the year.

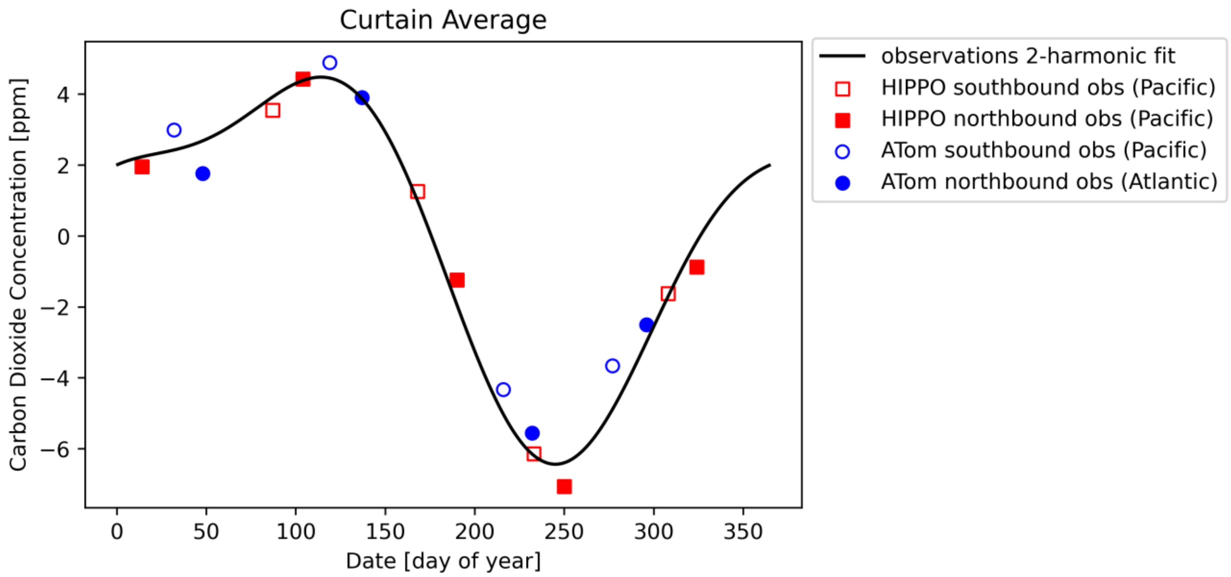


Figure 2.2 Two-harmonic fit to detrended average carbon dioxide concentration in ppm as a function of day of year for HIPPO and ATom flight campaigns in the atmospheric curtain between 20°N and 90°N in latitude and between 1000 hPa and 300 hPa in pressure. The points are found by filtering, detrending, interpolating and extrapolating to get a full altitude and latitude slice, then taking a pressure and cosine of latitude weighted average. The black line is the average of all fits to individual latitude-pressure bins with the annual mean removed. The direction of flight (southbound or northbound) is shown with filled and unfilled symbols because southbound flights occurred 2-3 weeks earlier than northbound flights.

We convert from a rate of change in dry air mole fraction to a rate of change in mass balance (MB) by multiplying the mole fraction by the mass of dry air north of 20 ° N and between the surface and 300 hPa in pressure. This mass is found using the ERA5 reanalysis fields used by Tracer Model 5 (TM5) within CarbonTracker 2019 (Krol et al., 2005; Jacobson et al., 2020). We use the time mean mass of $1.21087452 \times 10^{18}$ kg within the domain across the HIPPO and ATom time periods, neglecting annual and seasonal variations, which are less than 0.2%.

2.2.3 Flux Estimates Using Atmospheric Transport Models

Although the HIPPO and ATom observations over the remote ocean provide representative estimates of background values, the sampling and discretization method, zonal gradients, and mixing out of the domain result in differences between our MB time derivatives and zonal fluxes. Also, fossil fuel emissions and air-sea gas exchange make small contributions to the observed cycles. We use atmospheric transport models to account for the cumulative effect of: 1) atmospheric mixing across the southern boundary and above the pressure boundary; 2) spatial sampling biases associated with specific flight tracks; 3) zonal sampling bias; 4) temporal sampling biases associated with synoptic variability, subseasonal sample distribution, and interannual variability; and 5) contributions from fossil-fuel emissions and ocean uptake.

Atmospheric inversions provide optimal estimates of surface-atmosphere CO_2 exchange derived from both atmospheric CO_2 mole fraction data and initial estimates for land-atmosphere and ocean-atmosphere exchange in the context of an atmospheric transport model. We use posterior concentrations from three different inversions (Table 2.2) to reduce the uncertainty that may arise due to biases present in the choice of transport model as differences in transport have previously been shown to lead to large differences in optimized fluxes (Gurney et al., 2002; Schuh et al., 2019; Stephens et al., 2007).

CarbonTracker is a data assimilation system consisting of the TM5 atmospheric transport model coupled to an ensemble Kalman filter (Jacobson et al., 2020; Peters et al., 2007). TM5 is a global two-way nested transport model driven by 3-h meteorological forcing from the ERA5 operational forecast model (Krol et al., 2005). We used output from the most recent Carbon Tracker release (CT2019B, Jacobson et al., 2020), which includes optimized carbon fluxes through the HIPPO and ATom time period, and CO_2 mole fractions simulated along the flight paths for the HIPPO and ATom campaigns, which match the dates, times, and locations for the HIPPO and ATom data included in the GLOBALVIEWplus v5.0 ObsPack product (Cooperative Global Atmospheric Data Integration Project, 2019). CT2019B assimilated 460 time series datasets including data from the HIPPO and ATom campaigns. The datasets assimilated in CT2019B were mostly surface in situ, surface flask, and tower in situ observations from sites around the world.

The Model for Interdisciplinary Research on Climate version 4 atmospheric general circulation model based chemistry transport model (MIROC4-ACTM) provides posterior 4-D CO_2 fields and optimized surface fluxes through the HIPPO and ATom period (Chandra et al., 2022). Atmospheric CO_2 transport in MIROC4-ACTM is simulated by the Japan Agency for Marine-Earth Science and Technology's ACTM, a transport model driven by meteorological parameters from the Japanese 55-year Reanalysis (JRA55, Patra et al., 2018). We used the 2020 version of MIROC4-ACTM output and MIROC CO_2 mole fractions simulated along the flight paths for the HIPPO and ATom campaigns matching the dates, times, and locations for the HIPPO and ATom data included in the GLOBALVIEWplus v5.0 ObsPack product. MIROC4-ACTM assimilated surface flask data from 50 sites around the world (Chandra et al. 2022).

We used a third set of inverse modeling output from the Copernicus Atmosphere Monitoring Service (CAMS; Chevallier et al., 2005). Within CAMS, transport of atmospheric CO_2 is simulated by the global climate model of the Laboratoire de Météorologie Dynamique, zoom capacity (LMDZ) driven by meteorological parameters from ECMWF (Chevallier et al., 2005). We used CO_2 mole fractions simulated along the HIPPO and ATom flight paths matching the dates, times, and locations for the HIPPO and ATom data included in the GLOBALVIEWplus v5.0 ObsPack product, and posterior carbon fluxes from CAMS v20r1, which contains output through the HIPPO and ATom period. CAMS v20r1 assimilated surface air-sample data from 159 sites around the world.

Table 2.2 Inverse models used in this study.

	CT2019B	MIROC4-ACTM	CAMS
Years Available	2000-2018	1996-2018	1979-2020
Years Used	2009-2018	2009-2018	2009-2018
Transport	TM5	ACTM	LMDZ
Meteorology	ERA5	JRA55	ERA5
Resolution (lat x lon in degrees)	Globe 2x3, N America 1x1	Glb2.8x2.8	Glb1.9x3.75
Fossil Fuels	Miller and ODIAC	EDGARv432	GCP-GridFEDv2021.2
Reference	(Jacobson et al., 2020)	(Chandra et al., 2022)	(Chevallier et al., 2005)

For each model, we calculate the annual cycle of the northern extratropical net land flux by removing the long-term annual mean of the posterior land flux, which excludes fossil fuel emissions and ocean fluxes, from 2000 - 2018 at each grid cell and then taking an area-weighted average north of 20°N. We linearly interpolate between monthly means to get an annual cycle at daily resolution to allow for direct comparison to the aircraft-observation-derived MB.

To correct for bias, we match the model ObsPack output date, time, and location to the 10-second merge files for HIPPO and ATom, then repeat the analysis described in Section 2.2.2 using posterior CO_2 mole fractions simulated along the HIPPO and ATom flight tracks to calculate the curtain average and MB for each atmospheric inversion system. The averaged posterior land flux is then subtracted from the MB to derive a seasonal correction. The MB found using posterior CO_2 mole fractions simulated along the HIPPO and ATom flight tracks is the MB that would be observed if transport and fluxes in our world perfectly matched transport and fluxes in the model, with a time delay between the two curves due to the time takes the signal of land fluxes to reach the location where measurements are made. Thus, in each model, we assume the difference between the MB for each model (solid lines in Fig. 2.3a) and northern extratropical average posterior land flux (dashed lines in Fig. 2.3a) is due primarily to mixing outside the domain, model fluxes, and time delay, with additional influences listed above. We determine the correction for each model and subtract it from the observationally derived MB (dotted black line in Fig. 2.3b) resulting in transport-model specific flux estimates (solid color lines in Fig. 2.3b). The average difference across the three models (dashed black line in Fig. 2.3b) is subtracted from the observationally derived MB to estimate the seasonal cycle of the average net flux, hereafter referred to as the “flux cycle”, into the atmosphere (solid black line in Fig. 2.3b).

We find that the contribution of atmospheric transport uncertainty for the large spatial scale over which we average is small. In particular, noted variations in representations of vertical mixing (Stephens et al., 2007; Schuh et al. 2019) may change the distribution of CO_2 within our domain but not the domain average.

2.2.4 Growing Season Net Flux from Seasonal Flux Cycles

We then calculate the net atmospheric carbon exchange during the growing season, or GSNF, as the integral of the flux cycle during the growing season, defined to be when the bias corrected flux cycle is negative, which is nominally equivalent to the period when detrended atmospheric CO_2 is declining, primarily due to additional uptake by the biosphere as GPP outpaces respiration.

By detrending the observations, this estimate excludes the annual mean flux of CO_2, which itself includes fossil fuel emissions and terrestrial and oceanic sinks. Thus, our estimate of GSNF reflects the seasonal-only component of terrestrial exchange; the actual net uptake by the terrestrial biosphere during the growing season is larger when the annual component (long-term sink) is included. This approach is consistent with prior use of GSNF (e.g. Fung et al., 1983; Yang et al., 2007). Seasonal variations in fossil-fuel emissions and air-sea exchange contribute to seasonal variations in atmospheric CO_2, but these influences are small at about 3% and 5% of land exchange respectively on average for the three inversion systems, and have been removed by our use of the land flux in our model-based bias correction.

Earth System Models (CMIP5 and CMIP6)

The Coupled Model Intercomparison Project (CMIP) is an international, multi-model research intercomparison project whose purpose is to compare a coordinated set of simulations from ESMs in order to gain a better understanding of our ability to model climate change and associated feedbacks (Friedlingstein et al., 2006). The ESMs that participate in the CMIP simulate relevant physical, chemical, and biological processes within the coupled Earth system

(Eyring et al., 2016) using models developed by individual modeling centers worldwide, with the goal of including the most important processes that feed back into the climate system.

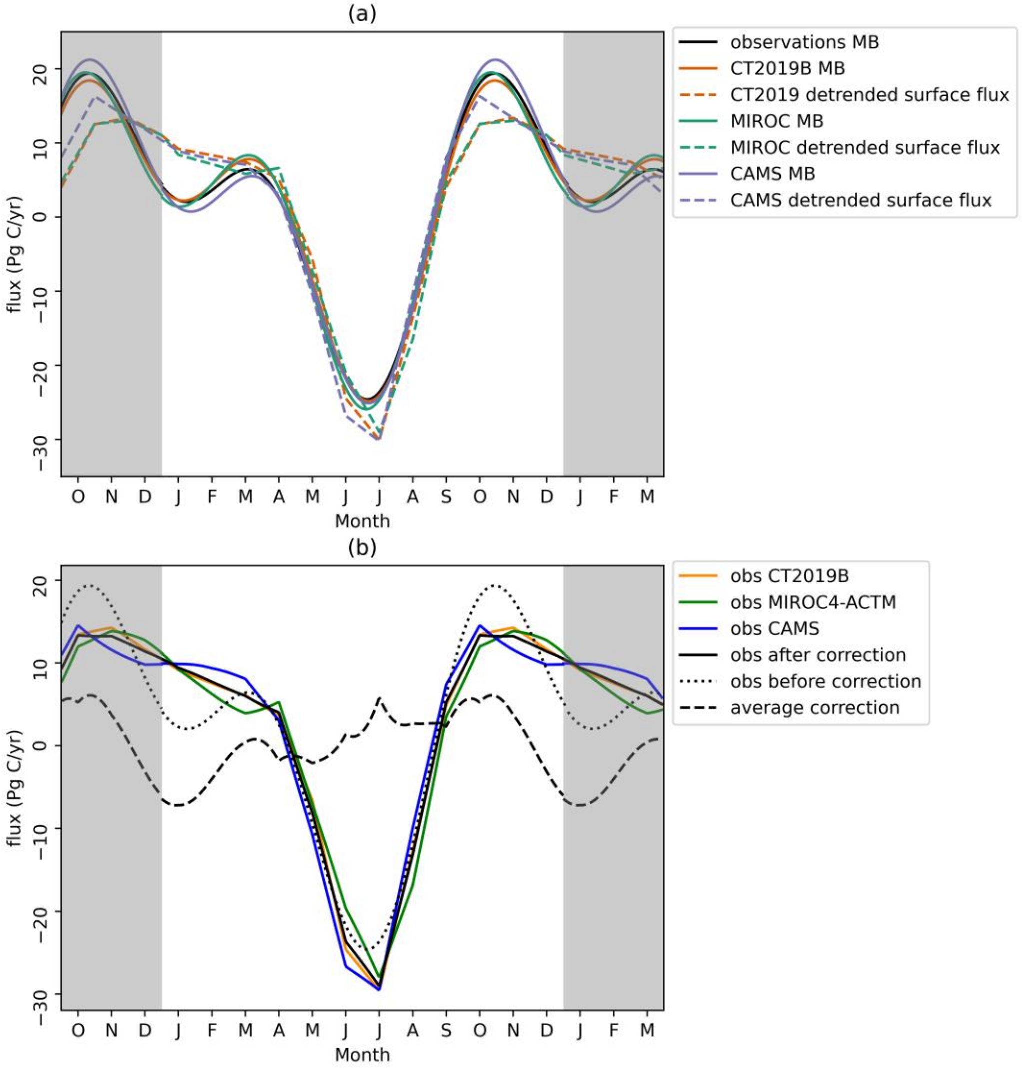


Figure 2.3 (a) Time derivative of concentrations from observations and inverse models, along with model fluxes. The dashed lines the area weighted average of posterior land fluxes from each inversion system in the domain 20°N-90°N. The solid lines are found by using the carbon dioxide mole-fractions along the flight track for each model to calculate the MB as described in Section 2.2. The solid black line is the time derivative of concentrations using the HIPPO and ATom observations. The estimated flux for the observations is bias corrected by finding the difference between the dotted and solid lines for a given model and applying that difference to the time derivative of the concentration. (b) Estimated flux after bias correction. The colored lines are

found by calibrating using only the model indicated and the solid black line is found using the average correction. The dotted black line is the time derivative of concentration before the correction.

Here, we analyzed the historical simulations for CMIP5 and CMIP6 (Table 3) that span the period from 1850 to 2005 for CMIP5 and to 2014 for CMIP6. We analyzed the CO_2-concentration driven historical simulations in which environmental forcing, such as greenhouse gas concentrations and solar forcing, are prescribed. Land and ocean fluxes are allowed to evolve prognostically in response to greenhouse gasses and other forcings (Eyring et al., 2016).

The northern extratropical seasonal land flux for CMIP5 and CMIP6 models is calculated by removing the long-term annual mean and taking the area-weighted average of each model's NBP output north of 20°N, linearly interpolating between monthly mean values, similar to the method used on the inversion posterior fluxes (Section 2.2.3). In addition, we fit a second-order harmonic to the NBP seasonal cycle to find seasonal timing within the models. We define the growing season in each model as the period for which the model simulates net terrestrial uptake, which allows us to evaluate each model's growing season start and end dates against those inferred from the aircraft observations. In ESMs, NBP reflects the balance of gross photosynthetic uptake, ecosystem respiration, and disturbance and harvest fluxes, and corresponds to the land-atmosphere carbon exchange, making it comparable with our observationally derived flux. We note that ESMs generally do not represent lateral carbon fluxes in rivers, but we expect these to have a relatively minor contribution to our observed seasonal variations. We average multiple years of NBP output from the CMIP models to derive a climatology; for CMIP6, we average over 2009 - 2014, and for CMIP5, we average over 2000-2005 because the output is not available through the HIPPO and ATom timeframe. We evaluate the ensembles as a whole by taking the median of each ensemble of models. We also evaluate area-weighted averages of three major component fluxes: GPP, RH, and NPP. Instead of analyzing the fluxes integrated over the growing season, we analyze the fluxes integrated over the entire year for GPP, RH, and NPP, which still correlate with the GSNF and are more useful than the seasonal fluxes when analyzing the carbon budget (e.g. Ballantyne et al., 2015; Tans et al., 1990). For the flux analysis, we use a subset of the models (9 of 12 CMIP5 models and 13 of 17 CMIP6 models, Table 2.3) for which historical NBP, GPP, RH, and NPP are available.

Table 2.3 Earth system models used in this study. Models in bold are included in the subset used to analyze GPP, RH, and NPP

Model	Generation	Land Component	Ocean Component	Institution ID	Reference
ACCESS-ESM1-5	**CMIP6**	**CABLE2.4**	**ACCESS-OM2 (MOM5)**	**CSIRO**	**Ziehn et al., 2019**
CanESM5	**CMIP6**	**CLASS3.6/CTEM1.2**	**NEMO3.4.1**	**CCCma**	**Swart et al., 2019**
CESM2	**CMIP6**	**CLM5**	**POP2**	**NCAR**	**Danabasoglu et al., 2019**
CESM2-FV2	CMIP6	CLM5	POP2	NCAR	Danabasoglu et al., 2019
CESM2-WACCM	**CMIP6**	**CLM5**	**POP2**	**NCAR**	**Danabasoglu et al., 2019**
CESM2-WACCM-FV2	CMIP6	CLM5	POP2	NCAR	Danabasoglu et al., 2019
CMCC-CM2-SR5	**CMIP6**	**CLM4.5**	**NEMO3.6**	**CMCC**	**Lovato et al., 2020**
CMCC-ESM2	**CMIP6**	**CLM4.5**	**NEMO3.6**	**CMCC**	**Lovato et al., 2021**
GISS-E2-1-G	CMIP6	GISS LSM	GISS	NASA-GISS	NASA/GISS 2018
GISS-E2-1-H	CMIP6	GISS LSM	HYCOM	NASA-GISS	NASA/GISS 2019
IPSL-CM6A-LR	**CMIP6**	**ORCHIDEE**	**NEMO-OPA**	**IPSL**	**Boucher et al., 2018**
MPI-ESM-1-2-HAM	**CMIP6**	**JSBACH3.2**	**MPIOM1.63**	**HAMMOZ-Consortium**	**Neubauer et al., 2019**
MPI-ESM1-2-LR	**CMIP6**	**JSBACH3.2**	**MPIOM1.63**	**MPI-M**	**Wieners et al., 2019**

NorCPM1	**CMIP6**	**CLM4**	**MICOM1.1**	**NCC**	**Bethke et al., 2019**
NorESM2-LM	**CMIP6**	**CLM**	**MICOM**	**NCC**	**Seland et al., 2019**
NorESM2-MM	**CMIP6**	**CLM**	**MICOM**	**NCC**	**Bentsen et al., 2019**
TaiESM1	**CMIP6**	**CLM4**	**POP2**	**AS-RCEC**	**Lee et al., 2020**
CanESM2	**CMIP5**	**CLASS2.7 and CTEM1**	**CanOM4 and CMOC1.2**	**CCCma**	**Chylek et al., 2011**
CCSM4	CMIP5	CLM4	POP2	NCAR	Gent et al., 2011
CESM1-BGC	CMIP5	CLM	BEC	NSF-DOE-NCAR	Long et al., 2013
GFDL-ESM2G	**CMIP5**	**LM3**	**TOPAZ**	**NOAA GFDL**	**Dunne et al., 2013**
HadGEM2-CC	**CMIP5**	**MOSES2 and TRIFFID**	**HadGOM2**	**MOHC**	**Collins et al., 2011**
HadGEM2-ES	**CMIP5**	**MOSES2 and TRIFFID**	**HadGOM2**	**MOHC**	**Collins et al., 2011**
INM-CM4.0	CMIP5	-	-	INM	Volodin et al., 2010
IPSL-CM5A-LR	**CMIP5**	**ORCHIDEE**	**ORCA2**	**IPSL**	**Dufresne et al., 2013**
IPSL-CM5A-MR	**CMIP5**	**ORCHIDEE**	**ORCA2**	**IPSL**	**Dufresne et al., 2013**
MIROC-ESM	**CMIP5**	**MATSIRO**	**COCO**	**MIROC**	**Watanabe et al., 2011**
MIROC-ESM-CHEM	**CMIP5**	**MATSIRO**	**COCO**	**MIROC**	**Watanabe et al., 2011**
NorESM1-M	**CMIP5**	**CLM**	**MICOM**	**NCC**	**Tjiputra et al., 2013**

2.3 Results

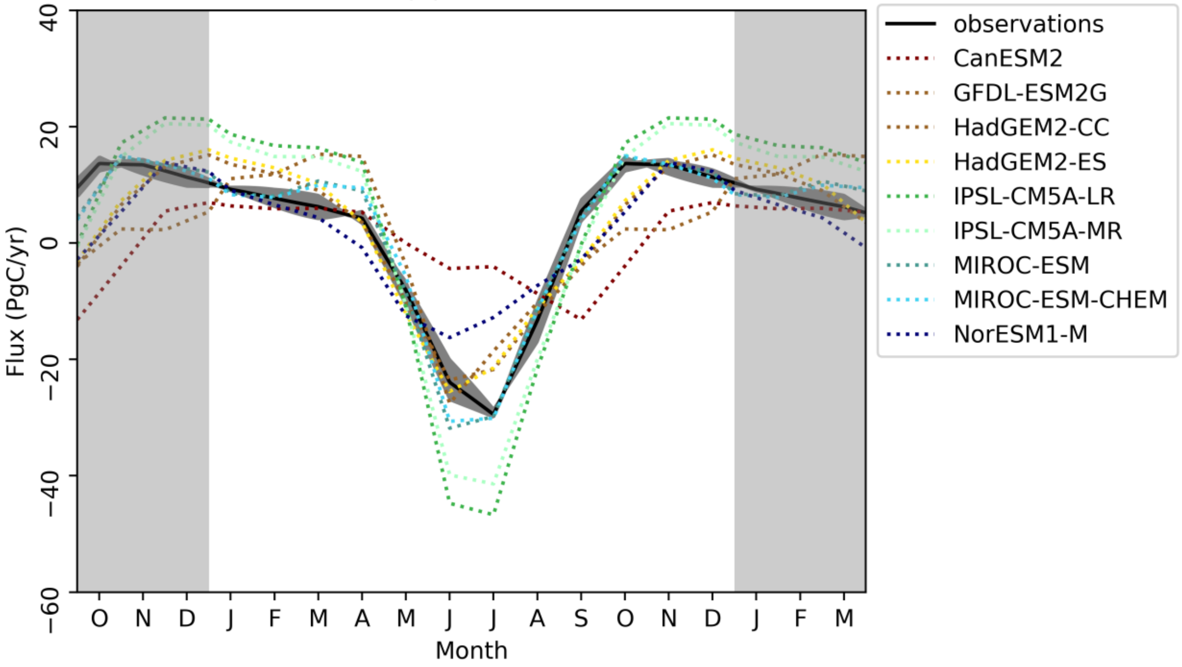
Observations of atmospheric carbon dioxide from the HIPPO and ATom aircraft campaigns are used to estimate a GSNF of 5.7 ± 0.3 Pg C out of the atmosphere north of 20°N averaged over the period 2009-2018 (Fig. 2.5). This value is equivalent to the net CO_2 exchange between the land and atmosphere during the growing season after removing the annual mean. The growing season is defined as the period when seasonal fluxes are negative (net uptake by land greater than the annual mean) and occurs between day 117 in late April and day 248 in early September. This corresponds to the day when the curtain average is maximum to the day when the curtain average is minimum (Fig. 2.2). The flux cycle shows maximum uptake on day 196 (Fig. 2.3b).

We conducted sensitivity tests to ensure that our choices for the latitudinal boundary of the domain, vertical boundary of the domain, and bin size did not have a large influence on the calculated GSNF. Expanding the region north (south) of 20°N by 5°N resulted in a decrease (increase) in the strength of the GSNF of just under 1%. The GSNF was also generally robust to the choice of pressure ceiling for the aircraft observations, increasing by just over 1% when we instead used 350 hPa as a ceiling. The relative standard deviation for all boundary combinations tested (all possible combinations of pressure cutoffs of 300 hPa, 325 hPa, 350 hPa, 375 hPa, and 400 hPa with latitude cutoffs of 20°N and 25°N) was 2%. Similarly, we saw just over a 1% increase in GSNF magnitude for a doubling of latitude or pressure bin size. The relative standard deviation for all bin sizes tested (all possible combinations of pressure bin sizes of 25 hPa, 50 hPa, and 100 hPa with latitude bin sizes of 5 ° and 10 °) was 1%.

Given the small differences that the choice of boundary and bin size make on the magnitude of the calculated GSNF, most of the uncertainty in GSNF results from the transport model bias correction process. We see a spread of 0.2 Pg C or 4% when using CT2019B alone to bias correct vs. using CAMS alone to bias correct. Additionally, the interannual variability, calculated as the standard deviation of the GSNF estimated from the average of the three inversion fluxes for each year over the period 2009-2018, is less than 0.2 Pg C or just over 3%. When adding the uncertainty from bin size, boundary choice, bias correction, and interannual variability in quadrature, assuming uncorrelated errors, the 1-sigma error on GSNF is 0.3 Pg C and on the start and end of the growing season is 2 days.

Considering the extensive altitude-latitude coverage of the aircraft observations, the inferred flux represents a unique and robust hemispheric estimate of terrestrial biosphere exchange and its seasonal phasing. We used the observationally inferred GSNF metric to evaluate NBP and its seasonal phasing from the CMIP5 and CMIP6 ensembles (Fig. 2.5). CMIP6 models (mean GSNF: 5.3 ± 1.6 Pg C, range: 1.5 - 6.9 Pg C) on the whole have less spread than CMIP5 models (mean GSNF: 5.7 ± 2.4 Pg C, range: 2.5 - 10.0 Pg C) (Fig 2.5). Three of the 13 CMIP6 ensemble members evaluated, CESM2, CESM2-WACCM, and MPI-ESM1-2-LR, were within 0.3 Pg C of the observed value, and most of the models with a large bias underestimated the GSNF.

(a) CMIP5



(b) CMIP6

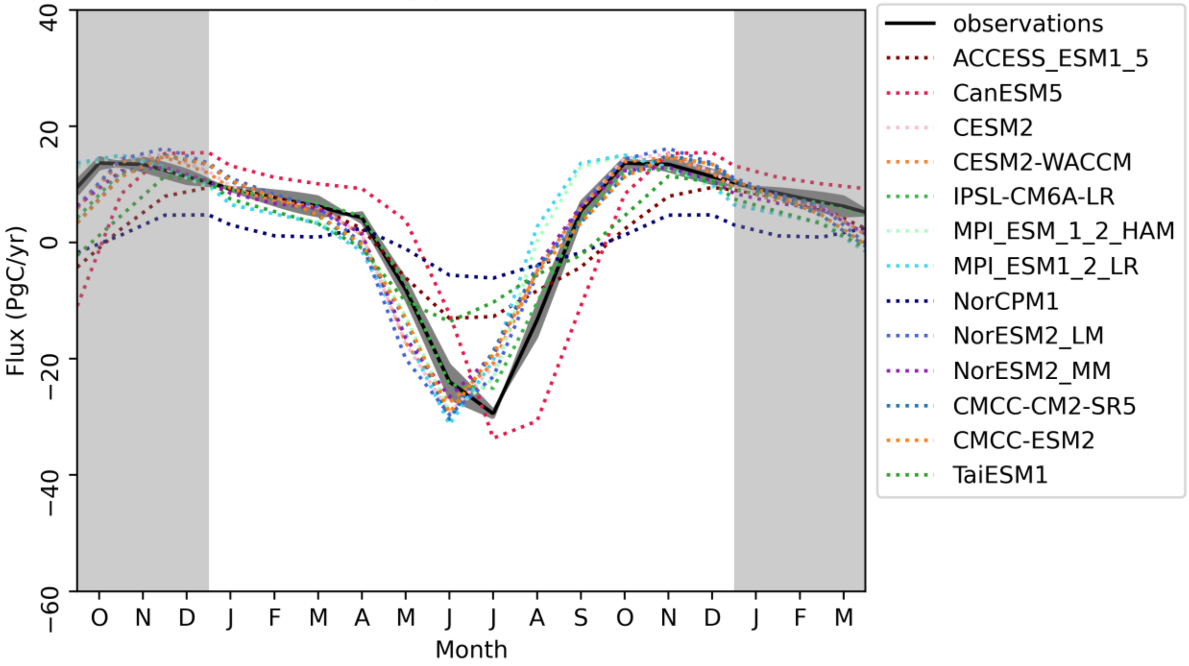


Figure 2.4 Corrected flux estimated from the HIPPO and ATom campaigns in comparison to area-weighted average NBP in the same domain from the (a) CMIP5 and (b) CMIP6 models. The bias corrected observation error is the standard deviation between correction using the three different inverse models. While the spread in magnitude and timing of the flux in CMIP6 models is smaller than that of CMIP5 models, there is still disagreement between models.

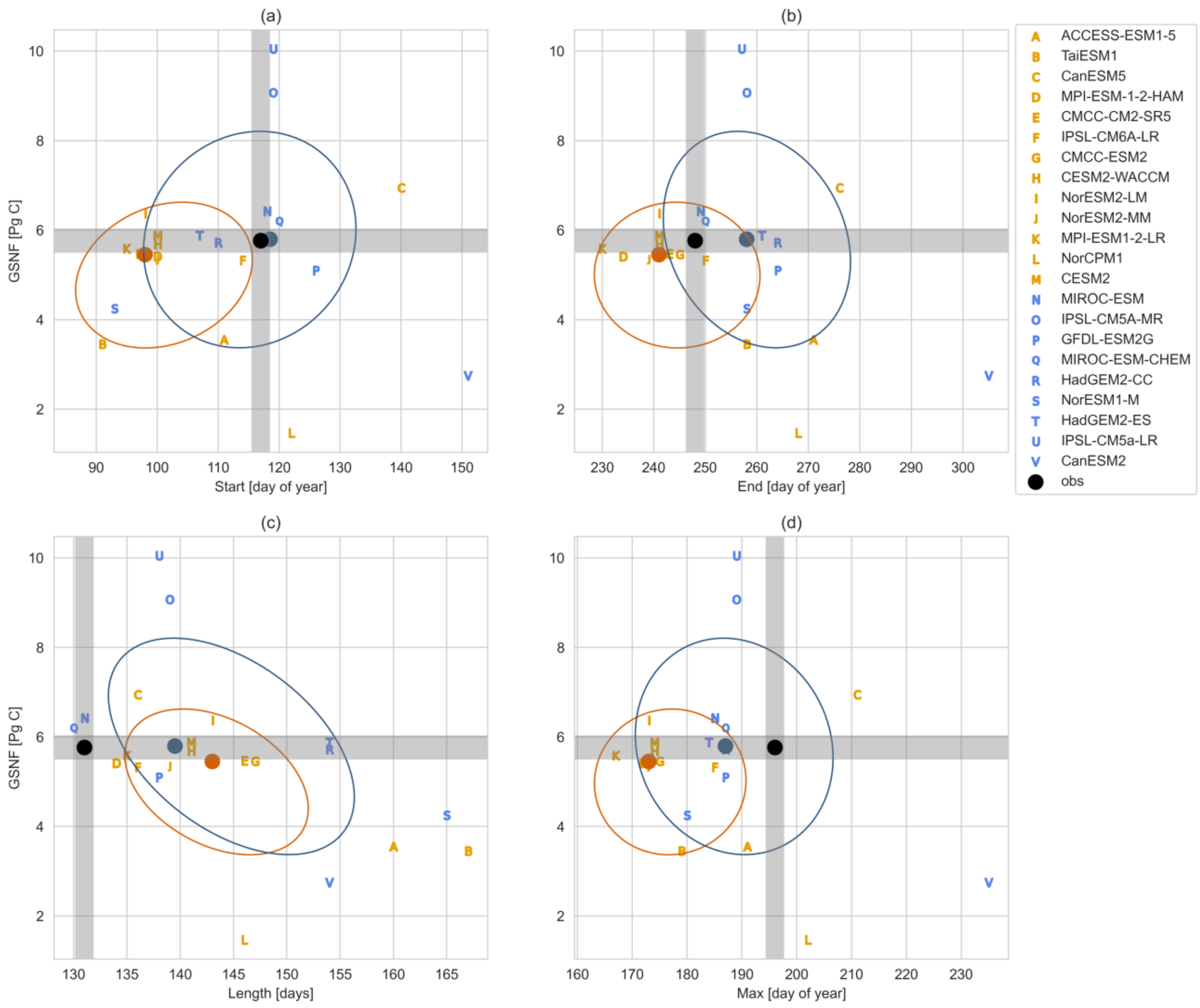


Figure 2.5 GSNF plotted against the (a) start of the season, defined to be the first day when the seasonal component of atmospheric CO2 is decreasing (seasonal component of flux changes from positive to negative), (b) end of the season, defined to be the last day when the seasonal component of atmospheric CO2 is decreasing (seasonal component of flux changes from negative to positive), (c) length of the season, and (d) max of season, defined to be the day when flux is most negative. The black point is the number inferred from the observations with the gray lines showing uncertainty. CMIP5 models are shown in blue and CMIP6 models are shown in orange. The blue and orange points are the multi-model mean for the CMIP5 and CMIP6 ensembles respectively. The surrounding ellipses show the covariance to one standard deviation. Only models where GPP, RH, and NPP output was available are included.

We note that some modeling centers showed substantial improvement in capturing GSNF between CMIP5 and CMIP6 (Fig. 2.6) by decreasing the absolute value of their z-score, which indicates how many standard deviations away from the observed value a model falls. For example, the two versions of the IPSL model overestimate GSNF by more than 4 Pg C in CMIP5, which improved substantially in CMIP6 to underestimate the GSNF by less than 1 Pg C. In both CMIP5 and CMIP6, the majority of ensemble members underestimate the seasonal flux, falling below the horizontal gray bar in Fig. 2.5.

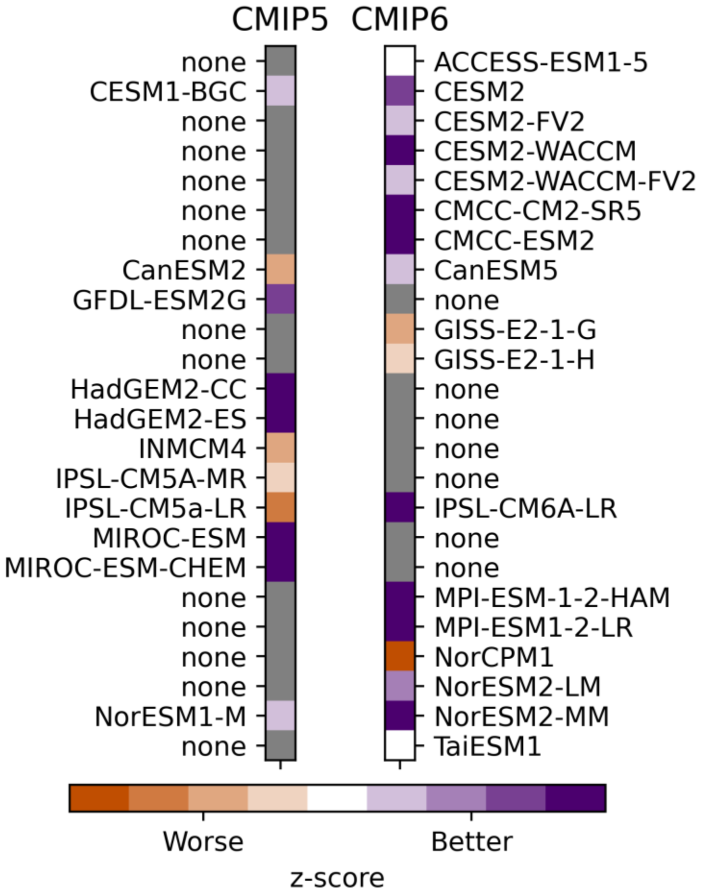


Figure 2.6 CMIP5 and CMIP6 model absolute value of z-score calculated for all models. The color gray and the label none has been used when one generation of a model is not used or not existent.

A decrease in GSNF model spread in the newer generation has not necessarily led to an improved agreement between models and observations on the phasing of the seasonal cycle. For example, The CMIP5 ensemble median start day (Julian day 119 ± 17) is closer to the observed start day (Julian day 117 ± 2) than the CMIP6 ensemble mean start day (Julian day 98 ± 14) (Fig. 2.5a). None of the 13 CMIP6 models evaluated fell within 2 days of the observed start day while three of the 9 CMIP5 models evaluated fell within the 2-day uncertainty range (Fig. 2.5a). The observations suggest that the growing season onset has become more biased in CMIP6. However, the CMIP6 ensemble did show a smaller bias than CMIP5 for the model-median end day, which was Julian day 241 ± 16 for CMIP6 and 258 ± 18 for CMIP5, in contrast to day 248 ± 2 in the observations (Fig. 2.5b). At 143 ± 9 days, the CMIP6 ensemble median growing season length is 12 days longer than the observed length of 131 ± 2. In comparison, the CMIP5 ensemble median growing season length is 11 days longer at 140 ± 12 days (Fig. 2.5c). While the CMIP6 ensemble median end day compares more favorably with the observed end day than does the CMIP5 ensemble median end day, in both cases, the simulated growing season is longer than what is observed (Fig. 2.5c).

Seasonal phasing in general does not appear to be correlated with GSNF in the model ensembles, suggesting that phasing is not a dominant driver of GSNF spread among models. No correlation was seen ($r2 < 0.2$, $p > 0.05$) between GSNF magnitude and the start day, end day, length, or max day across models (Fig. 2.5). This suggests that factors other than the phasing of the growing season may explain inter-model differences.

We analyzed GPP, RH, and NPP for the subset of CMIP5 and CMIP6 ensemble members which include these outputs to see if these component fluxes might explain model disagreement on GSNF (Fig. 2.7). We found that in both the CMIP5 and CMIP6 ensembles, GSNF was correlated ($r^2 = 0.68$, $p < 0.05$) with GPP, where models with larger GPP generally had larger GSNF. Models with large GPP also tend toward higher respiration values with an r^2 value of 0.77 between GPP and RH, as GPP provides the inputs to support RH. RH showed a weaker correlation with GSNF than did GPP, however, the correlation is still moderately strong ($r^2 = 0.52$, $p < 0.05$). As expected, models with higher GPP values also tend toward higher NPP values, and the correlation between NPP and GSNF is moderately strong with an r^2 value of 0.66 and a p-value less than 0.05.

We see a large range for GPP, RH, and NPP across the CMIP5 and CMIP6 ensembles. The range in GPP is smaller for the CMIP6 ensemble (median = 51 ± 9 Pg C, spread = 37 Pg C) than for the CMIP5 ensemble (median = 60 ± 17 Pg C, spread = 47 Pg C) where the error on the median is one standard deviation. However, this decrease in the model range between ensemble generations is not seen for RH or NPP. The spreads for RH were 23 Pg C (median = 27 + 8 Pg C) for CMIP5 and 27 Pg C (median = 21 ± 7 Pg C) for CMIP6. NPP from CMIP5 and CMIP6 had median values of 32 ± 9 Pg C and 24 ± 7 Pg C respectively and a spread of 28 Pg C for both CMIP5 and CMIP6. We note that two models in the CMIP6 ensemble have GSNF values consistent with the observational constraint, MPI-ESM1-2-LR and IPSL-CM6A-LR, but the GPP values spanned by these models are over 10 Pg C, the RH values spanned by these models are 10 Pg C, and the NPP spanned by these models are 9 Pg C.

The strong correlations and large ensemble spread enabled us to indirectly constrain northern extratropical GPP, RH, and NPP through an “emergent constraint” (EC) approach (e.g. Eyring et al., 2019, Williamson et al., 2021; Simpson et al., 2021). ECs are correlations between some observable element X that varies across the ESM ensemble and some important variable Y assuming a physically meaningful relationship exists between X and Y. Here, we assumed GSNF to be X and assumed a physically meaningful relationship between the net flux and its component fluxes GPP, RH, and NPP (Fig. 2.7). The CMIP5 and CMIP6 ensembles can be analyzed separately to estimate GPP, RH, and NPP. The CMIP5 ensemble EC estimate for GPP is 58 ± 23 Pg C, for RH is 28 ± 13 Pg C, and for NPP is 31 ± 12 Pg C where the error is the 95% prediction interval from a hypothetical sample generated using a Monte Carlo simulation of n = 10000 and assuming a Gaussian distribution (Williamson and Sansom 2019). Similarly, the CMIP6 ensemble EC estimate for GPP is 54 ± 13 Pg C, for RH is 24 ± 13 Pg C, and for NPP is 26 ± 10 Pg C. However, we can also combine the two ensembles to create a larger sample size as the estimates agree within error and the modeled relationship does not change substantially between the two generations. The combined EC estimate for GPP is 56 ± 17 Pg C, for RH is 25 ± 13 Pg C, and for NPP is 28 ± 11 Pg C. Although we did not constrain autotrophic respiration (RA) explicitly, the constraint on GPP and NPP imply RA is 28 ± 20 Pg C where the error has been propagated by summing the errors in quadrature. Applying an emergent constraint to RA explicitly would likely give smaller errors.

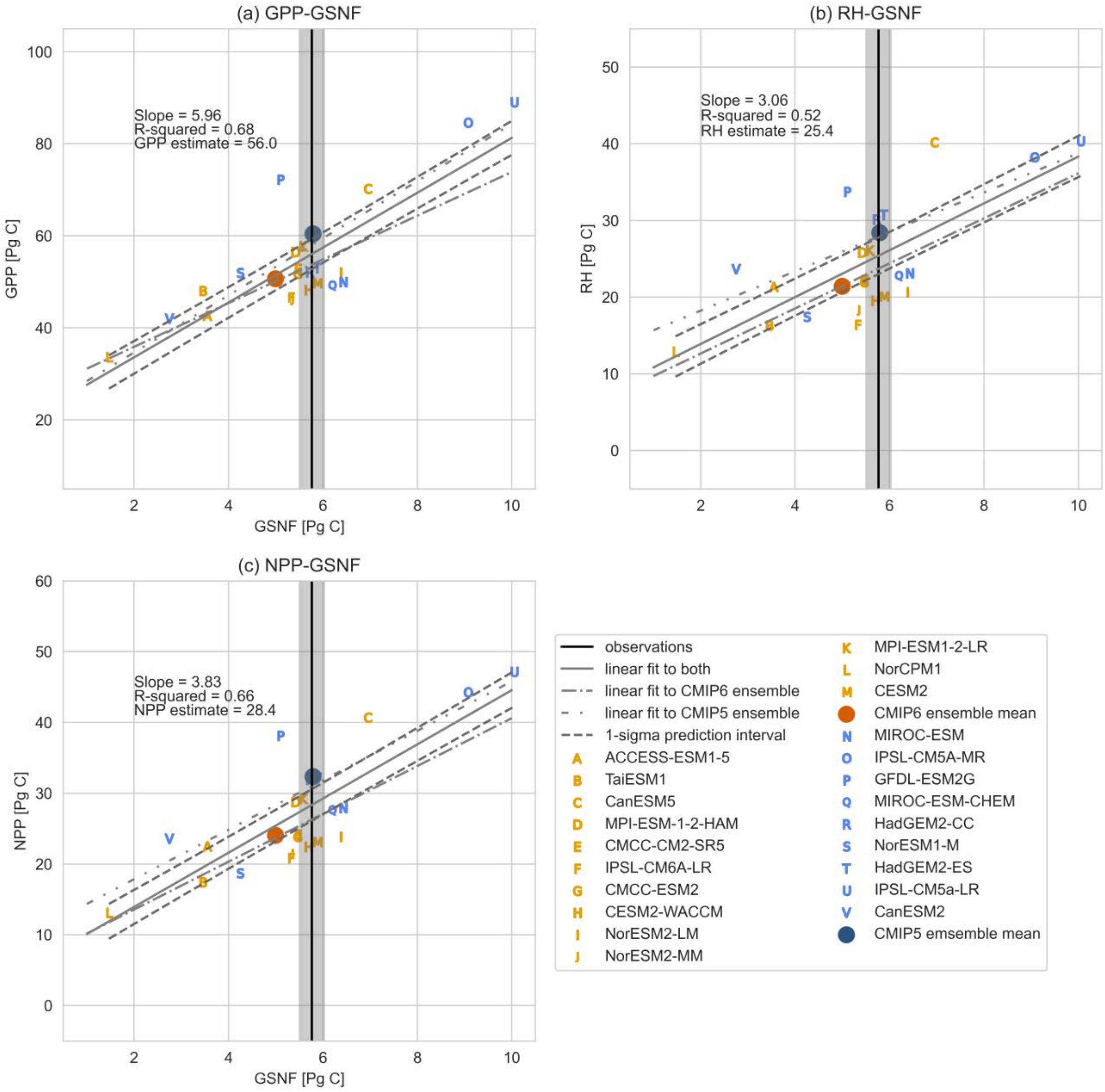


Figure 2.7 GSNF plotted as a function of (a) integrated GPP and (b) integrated RH, and (c) integrated NPP for the CMIP5 and CMIP6 models. The estimated GSNF from the HIPPO and ATom observations is shown in gray, CMIP5 models are shown in blue, and CMIP6 models are shown in orange. Only models where GPP, RH, and NPP data was available are included.

2.4 Discussion

Here we derive an observational constraint on northern extratropical GSNF from two novel aircraft campaigns that measured the atmospheric CO_2 curtain over remote oceans. Our estimated GSNF of 5.7 ± 0.3 Pg C is significantly smaller in magnitude than the GSNF of 7.9 Pg C yr^{-1} out of the atmosphere estimated north of 30°N by Yang et al. (2007) based on total column observations and spatially sparse aircraft profiles in North America. These differences are most likely due to differences in methodology, as Yang et al. used a combination of total column and aircraft measurements at 8 locations between 30°N and 70°N to scale fluxes from a terrestrial ecosystem model (TEM), while we used aircraft data that sampled nearly continuously between our cutoff of 20°N and roughly 87°N to estimate GSNF without using a specific TEM. When our cutoff is changed from 20°N to 30°N, our estimate increases to 5.8 Pg C. Fung et al. (1983) quantified GSNF in the Northern Hemisphere using a three-dimensional tracer transport model. When our cutoff is changed to the Equator, our estimate increases to 6.1 Pg C, which falls within the 3.4 - 10.7 Pg C range given by Fung et al. (1983). However, the range given by Fung et al. (1983) is large. The uncertainty on northern hemisphere GSNF could likely be reduced with application of multiple transport models to existing records.

We may expect to see an increase in GSNF over time due to the observed increase in the CO_2 seasonal cycle amplitude (SCA). Graven et al. (2013) saw an increase of 32-59 % in the NH seasonal cycle amplitude in the 50-year period from 1958-1963 to 2009-2011. This equates to an increase of 0.56 - 0.93 % yr^{-1}. If this trend continued through the HIPPO and ATom time period, we would expect to see a 4.0 - 6.7% increase across the 7 years between these two missions. Applying the methodology to HIPPO alone, which covers 2009-2011, gives a value of 11.2 ppm for the SCA of the curtain average versus 10.7 ppm when using ATom alone, which covers

2016-2018. This corresponds to a more than 1% decrease, contrary to the increase expected from Graven. However, this may not be accurate as there may not be enough data in either mission alone to fully constrain the seasonal cycle (Fig. S3) and account for interannual variability in CO_2 fluxes, which were shown by Jin et al. (2021) to be non-negligible.

Despite any interannual variability, the expected increase in GSNF is not apparent when comparing the flight-based estimates to previous estimates. Yang et al. (2007) estimate a larger GSNF despite covering an earlier time period. The large range given by Fung et al. (1983) means that the increase in GSNF may be seen if the true GSNF in 1982 fell within the range of 4.6 - 5.2 Pg C. Looking at surface records, Keeling and Graven (2021) saw an increase in the amplitudes at Mauna Loa and Barrow between 2010 and 2017 equating to 10.9% at Mauna Loa and 1.1% at Barrow using 5-year running means, suggesting possibly a shift to less SCA growth at high latitudes. These results should be interpreted with caution as large interannual variability may dominate when looking at the short time period of 2010-2017.

The HIPPO and ATom observations reflect the atmospheric mass of carbon integrated over a large latitudinal and altitudinal extent, thus, our estimate is less sensitive to specific representations of atmospheric transport than are other estimates. For example, the GSNF estimated over our domain, bias corrected with the three inverse models individually, only varied by just 0.2 Pg C or less than 4% (Fig. 2.3b), even while previous research has reported the relative spread of 10 simulations for the posterior annual mean northern extratropical land flux to be 13 % (Gaubert et al., 2019). Additionally, inverse models seem to be converging on the land flux as the GSNF estimated from the posterior inversion land fluxes averaged 5.8 Pg C, slightly larger than the observationally based estimate of 5.7 Pg C, and varied by just 0.2 Pg C. The small differences among the inverse models at the hemispheric scale mean that, even though it represented the largest source of uncertainty, the bias correction process imparted minimal uncertainty on our ultimate GSNF value.

Use of the transformed coordinate, $M_{\theta e}$, introduced by Jin et al. (2021) may further reduce uncertainty. $M_{\theta e}$ is defined as the dry air mass under a given equivalent potential temperature surface within a hemisphere and its relationship with latitude and altitude is nearly fixed over the seasonal cycle. When integrating to find the curtain average, Jin et al. (2021) show that using the transformed coordinate as an alternative to latitude reduces error in the curtain average due to sparse sampling and synoptic variability.

The low uncertainty on the hemispheric integral makes GSNF and its phasing robust targets for evaluating TEMs and the land components of ESMs. Although direct comparison with atmospheric CO_2 mole fraction has been used to evaluate ESMs previously, these comparisons rely on simulation of the three-dimensional atmospheric CO_2 mole fraction within the ESM (e.g., Keppel-Aleks et al., 2013) or require estimating CO_2 using an offline transport model or operator (e.g., Liptak et al, 2017). In contrast, our GSNF constraint can be used to evaluate ESMs at the flux level rather than relying on comparisons at the concentration level. Compared to our metric, the CMIP6 ensemble has shown some improvements relative to the CMIP5 ensemble, namely a reduction in the spread between models (5.5 Pg C vs. 7.3 Pg C) and a more favorable simulation of the timing of the end of the growing season (7 days early vs. 10 days late) when considering the ensemble median values (Fig. 2.5 b, c). However, there are still large disagreements, and in some ways, CMIP5 models perform better in relation to the observationally inferred flux than do CMIP6 models. For example, CMIP5 models outperform CMIP6 models in simulating the start of the growing season (2 days late vs. 19 days early) and the GSNF magnitude (5.7 ± 2.4 Pg C vs 5.3 ± 1.6 Pg C compared to observational value of 5.7 ± 0.3 Pg C) when considering the ensemble median values (Fig. 2.5). The models tend to underestimate the magnitude of GSNF on the whole with 3 of the 9 CMIP5 models and 8 of the 13 CMIP6 models underestimating the flux (falling below the horizontal gray lines in Figure 2.5), only 2 CMIP5 and 3 CMIP6 models falling within the uncertainty range, and 4 CMIP5 models and 2 CMIP6 models overestimating the net flux. This is consistent with previous results from Keppel-Aleks et al. (2013) suggesting that the Community Land Model version 4 (CLM4), a version of which is used as the land model for over one-third of the ESMs evaluated, underestimated GSNF. Three of the models with the largest underestimate of GSNF (NorCPM1, NorESM1-M, and TaiESM1) utilize CLM4 as their land model; however, we note that two of the best performing models (CESM2 and CESM2-WACCM) both utilize the Community Land Model version 5 (CLM5) as their land component, reflecting major improvement in CLM5 compared to CLM4 in simulating CO_2 seasonality (Lawrence et al., 2019). Correlations between GSNF and seasonal phasing were weak or nonexistent, meaning that the improvements in CLM5 are more likely due to improvements in simulations of the overall magnitude of photosynthesis and respiration than to improvements in simulations of the timing of the growing season.

The methods chosen to define the growing season have some effect on model-observation comparisons. We could have chosen to prescribe the growing season dates for calculating model metrics from the observations rather than allowing the dates to vary from model to model. The primary difference in the results using this approach is a slight reduction in the magnitude of the GSNF for most of the CMIP5 and CMIP6 models, a less strong correlation between GSNF and GPP, RH, and NPP, and a slight increase in the inferred values of northern GPP, RH, and NPP (Fig S2).

Although the aircraft CO_2 data provide a constraint on GSNF, this benchmark is also useful for constraining component fluxes that are difficult to infer observationally at large spatial scales. We examine the GPP, RH, and NPP of the ESM ensembles in relation to GSNF. GPP, RH, and NPP are components of the net flux, and we see moderate correlations between simulated GSNF and simulated productivity (both GPP and NPP) and RH. It is noteworthy that correlations between GSNF and the various annual fluxes are stronger than those between GSNF and seasonal timing, which may indicate that the magnitudes of photosynthesis and respiration are more dominant drivers of GSNF than seasonal phasing is (e.g., Valentini et al., 2000, Baldocchi et al., 2017). Cadule et al. (2010) analyzed three CMIP models and concluded that models generally underestimate the seasonal amplitude due to shortcomings in vegetation phenology and heterotrophic respiration response to climate. Our results generally support this role for discrepancies in component fluxes as driving a discrepancy in the resulting GSNF.

These correlations between GSNF and its component fluxes provide an opportunity to constrain GPP, RH, and NPP using an emergent constraint approach. Previous estimates of northern extratropical GPP are highly uncertain, with large disagreements between estimates. For example, Mao et al., (2012) used the Moderate Resolution Imaging Spectroradiometer (MODIS) GPP product to estimate Northern Hemisphere GPP averaged over 2000-2009 as 64.75 ± 0.97 Pg C yr^{-1}. In contrast, the FLUXCOM product, based on upscaled observations from FLUXNET sites using various machine learning approaches (Jung et al., 2019), estimated GPP north of 20°N averaged over 2009-2014 to be 48 Pg C yr^{-1}. Our GSNF-constrained value of 56 ± 15 Pg C for GPP is consistent with both of these estimates, given the larger error bars on the emergent constraint approach, and indicates that GPP falls between these two estimates. The large value for GPP relative to FLUXCOM is interesting in light of the uncertainty in global GPP. The range in global mean GPP magnitudes for 2008–2010 from FLUXCOM members is 106 - 130 Pg C yr^{-1} (Jung et al., 2020). This range covers the observation-based estimate of global mean GPP of 123 ± 8 Pg C yr^{-1}) found using eddy covariance flux data and various diagnostic models (Beer et al., 2010) but is smaller than the GPP magnitudes of 150–175 Pg C yr^{-1} derived from an isotope-based study (Welp et al., 2011). Our results may indicate a higher global mean GPP than the flux tower upscaling yields; however, we are constraining extratropical GPP which may be affected by different factors than tropical GPP. These results therefore need to be interpreted with caution.

2.5 Conclusions

Earth system models disagree in their simulation of large-scale carbon fluxes, making it crucial to evaluate models to contextualize their climate predictions. We have presented an approach to constrain the seasonal land flux using aircraft data from the HIPPO and ATom flight campaigns. The northern extratropical GSNF of 5.7 ± 0.3 Pg C derived from these observations can be used to evaluate prognostic model fluxes for net flux, directly, and for component fluxes via an emergent constraint approach. We note that this constraint is tied to atmospheric transport models because the flux is bias corrected by comparing posterior mole fractions from inverse models with their fluxes. We found, however, that at the hemispheric scale, the constraint is robust to the choice of atmospheric transport model since common transport errors tend to cancel out at this scale.

When compared to CMIP5 and CMIP6 models, the inferred GSNF suggests a larger net flux and shorter growing season than simulated in both model ensembles. This gives modelers an additional observation to target during model development and could be added to a benchmarking system such as the International Land Model Benchmarking (ILAMB) System

(Collier et al., 2018). While there is decreased model spread between CMIP6 models, this benchmark also highlights some of the ways in which CMIP6 models have not improved from CMIP5, such as in simulating the start of the growing season.

Correlations within the CMIP5 and CMIP6 ensembles allowed us to apply an emergent constraint approach to estimate northern hemisphere GPP, RH, and NPP. We found that the GSNF-constrained value for GPP is at the higher end of a range of estimates from FLUXCOM, a commonly used ensemble of data products of upscaled biosphere-atmosphere fluxes.

Overall, the HIPPO and ATom inferred GSNF provides a robust metric that allows for the evaluation of large-scale fluxes in flux space and sheds light on component fluxes, filling a need highlighted by Collier et al. (2018) for land model benchmarking. More regular global scale airborne tomography could resolve GSNF at higher time resolution and leverage carbon cycle interannual variability for improved tests of ESM process representations.

Acknowledgments

This research was partially supported by the Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Science Focus Area, which is sponsored by the Regional and Global Model Analysis (RGMA) activity of the Earth & Environmental Systems Modeling (EESM) Program in the Earth and Environmental Systems Sciences Division (EESSD) of the Office of Biological and Environmental Research (BER) in the US Department of Energy Office of Science.

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977. The HIPPO data were collected using NSF's Lower Atmosphere Observing Facilities, which are managed and operated by NCAR's Earth Observing Laboratory.

We thank the HIPPO science team and the flight crew and support staff of the NCAR Research Aviation Facility. We thank the ATom science team and the flight crew and support staff of the NASA DC-8, which is supported by the NASA Airborne Science Program and Earth Science Project Office. For additional support of HIPPO and ATom airborne CO_2 and N_2O measurements, we thank S. Afshar, J. Bent, J. Budney, B. Daube, G. Diskin, J. Elkins, Y. Gonzalo Ramos, J. Higgs, E. Hintsa, R. Jimenez, E. Kort, P. Lang, T. Legard, E. Moglia, S. Montzka, F. Moore, D. Neff, T. Newberger, B. Paplawsky, S. Park, J. Pittman, G. Santoni , A. Watt, S. Wofsy, S. Wolter, and B. Xiang.

We would like to thank the climate modeling groups (listed in Table 2 and 3 of this paper) for producing and making freely available their model output.

PKP is partly supported by the Arctic Challenge for Sustainability II grant (JPMXD1420318865) of the Ministry of Education, Science, Culture and Sports (MEXT), Japan.

Open Research

HIPPO data are available at the Earth Observing Laboratory (EOL) HIPPO data archive at http://www.eol.ucar.edu/projects/hippo/. (http://doi.org/10.3334/CDIAC/HIPPO_010).

ATom data are available at the Oak Ridge National Laboratory DAAC at https://daa.ornl.gov/ATOM/campaign/.(https://doi.org/10.3334/ORNLDAAC/1925).

CarbonTracker CT2019B results provided by NOAA ESRL, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov.(https://doi.org/10.25925/20201008)

MIROC4-ACTM inversion results are available on Zenodo: https://doi.org/10.5281/zenodo.5776197

CAMS inversion results provided by the Copernicus Atmoshere Monitoring Service, part of the European Union’s Earth observation programme, from the website at https://ads.atmosphere.copernicus.eu/.CMIP5 and CMIP6 model output files are freely available from the Earth System Grid Federation data replication centers and used under creative commons license CC BY-SA 4.0 https://creativecommons.org/licenses/by-sa/4.0/.

2.6 References

Akima H., "A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points", ACM Transactions on Mathematical Software, Vol. 4, No. 2, June 1978, pp. 148-159. Copyright 1978, Association for Computing Machinery, Inc., reprinted by permission.

Alexandrov, G. A. (2014). Explaining the seasonal cycle of the globally averaged CO2 with a carbon-cycle model. *Earth System Dynamics*, 5(2). https://doi.org/10.5194/esd-5-345-2014

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V.,et al.. (2020). Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences*, *17*(16), 4173–4222. https://doi.org/10.5194/bg-17-4173-2020

Baldocchi, D., Chu, H., & Reichstein, M. (2018). Inter-annual variability of net and gross ecosystem carbon fluxes: A review. Agricultural and Forest Meteorology, 249, 520–533. https://doi.org/10.1016/j.agrformet.2017.05.015

Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P., & White, J. W. C. (2012). Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. Nature. https://doi.org/10.1038/nature11299

Ballantyne, A. P., Andres, R., Houghton, R., Stocker, B. D., Wanninkhof, R., et al.(2015). Audit of the global carbon budget: estimate errors and their impact on uptake uncertainty. *Biogeosciences*, *12*, 2565–2584. https://doi.org/10.5194/bg-12-2565-2015

Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., et al.(2010). Terrestrial gross carbon dioxide uptake: Global distribution and covariation with climate. *Science*, *329*(5993), 834–838. https://doi.org/10.1126/science.1184984

Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., Hinton, T. J., et al. (2011). The HadGEM2 family of Met Office Unified Model climate configurations. Geoscientific Model Development, 4(3), 723–757. https://doi.org/10.5194/gmd-4-723-2011

Bent, J. (2014) Airborne oxygen measurements over the Southern Ocean as an integrated constraint of seasonal biogeochemical processes, PhD thesis, University of California, San Diego, USA, 2014.

Bentsen, Mats; Oliviè, Dirk Jan Leo; Seland, Øyvind; Toniazzo, Thomas; Gjermundsen, Adaet al. (2019). NCC NorESM2-MM model output prepared for CMIP6 CMIP historical.

Version 20191126 Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.8040 .

Bethke, Ingo; Wang, Yiguo; Counillon, François; Kimmritz, Madlen; Fransner, Filippa et al. (2019). NCC NorCPM1 model output prepared for CMIP6 CMIP historical.*Version 20190921* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.10894 .

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. Journal of Advances in Modeling Earth Systems, 12(7). https://doi.org/10.1029/2019MS002010

Boucher, Olivier; Denvil, Sébastien; Levavasseur, Guillaume; Cozic, Anne; Caubel, Arnaud; et al.(2018). IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP historical. *Version 20180711* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.5195 .

Cadule, P., Friedlingstein, P., Bopp, L., Sitch, S., Jones, C. D., Ciais, P., et al. (2010). Benchmarking coupled climate-carbon models against long-term atmospheric CO2 measurements. Global Biogeochemical Cycles, 24(2). https://doi.org/10.1029/2009GB003556

Canadell, J. G., P. M. S. Monteiro, et al. (2021), Global Carbon and other Biogeochemical Cycles and Feedbacks., in Climate Change 2021: The Physical Science Basis. Contribution of WG I to the Sixth Assessment Report of the IPCC, edited by L. G. Masson-Delmotte, et al., Cambridge Univ. Press.

Chandra, N., Patra, P. K., Niwa, Y., Ito, A., Iida, Y., et al. (2022).Estimated regional CO_2 flux and uncertainty based on an ensemble of atmospheric CO_2 inversions, Atmos. Chem. Phys., in press, 2022.

Cherchi, A., Fogli, P. G., Lovato, T., Peano, D., Iovino, D., Gualdi, S., et al. (2018). Global mean climate and main patterns of variability in the CMCC-CM2 coupled model. Journal of Advances in Modeling Earth Systems, 2018MS001369. https://doi.org/10.1029/2018MS001369

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F. M., et al. (2005). Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. Journal of Geophysical Research Atmospheres, 110(24), 1–13. https://doi.org/10.1029/2005JD006390

Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., & Ramonet, M. (2010). Atmospheric inversions for estimating CO_2 fluxes: Methods and perspectives. Climatic Change. https://doi.org/10.1007/s10584-010-9909-3

Cleveland, R.B. and Cleveland, W.S. (1990) STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6, 3-33.

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., et al. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. Journal of Advances in Modeling Earth Systems. https://doi.org/10.1029/2018MS001354

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 [Data set]. NOAA Earth System R

Danabasoglu, G., Lamarque, J. -F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). Journal of Advances in Modeling Earth Systems, 12(2). https://doi.org/10.1029/2019MS001916

Danabasoglu, Gokhan (2019). NCAR CESM2 model output prepared for CMIP6 CMIP historical. *Version 20190116* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.7627

Danabasoglu, Gokhan (2019). NCAR CESM2-FV2 model output prepared for CMIP6 CMIP historical. *Version 20190731* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.11297 .

Danabasoglu, Gokhan (2019). NCAR CESM2-WACCM model output prepared for CMIP6 CMIP historical. *Version 20190131* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.10071 .

Danabasoglu, Gokhan (2019). NCAR CESM2-WACCM-FV2 model output prepared for CMIP6 CMIP historical. *Version 20190831* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.11298 .

Daube, B. C., Boering, K. A., Andrews, A. E., & Wofsy, S. C. (2002). A High-Precision Fast-Response Airborne CO_2 Analyzer for In Situ Sampling from the Surface to the Middle Stratosphere. Journal of Atmospheric and Oceanic Technology, 19(10). https://doi.org/10.1175/1520-0426(2002)019<1532:AHPFRA>2.0.CO;2

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., et al. (2013). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. Climate Dynamics, 40(9–10), 2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., et al. (2013). GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics*. Journal of Climate, 26(7), 2247–2267. https://doi.org/10.1175/JCLI-D-12-00150.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Cox, P. M., Flato, G. M., Gleckler, P. J., Abramowitz, G., Caldwell, P., et al. (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, *9*(2), 102–110. https://doi.org/10.1038/s41558-018-0355-y

Fernández-Martínez, M., Sardans, J., Chevallier, F., Ciais, P., Obersteiner, M., Vicca, S., et al. (2019). Global trends in carbon sinks and their relationships with CO_2 and temperature. Nature Climate Change, 9(1). https://doi.org/10.1038/s41558-018-0367-7

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., von Bloh, W., Brovkin, V., et al. (2006). Climate-carbon cycle feedback analysis: Results from the C4MIP model intercomparison. Journal of Climate, 19(14), 3337–3353. https://doi.org/10.1175/JCLI3800.1

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

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., et al. (2022). Global Carbon Budget 2021. *Earth System Science Data*, *14*(4), 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Friend, A. D., Arneth, A., Kiang, N. Y., Lomas, M., Ogée, J., Rödenbeck, C., et al. (2007). FLUXNET and modelling the global carbon cycle. Global Change Biology. https://doi.org/10.1111/j.1365-2486.2006.01223.x

Fung, I. Y., Doney, S. C., Lindsay, K., & John, J. (2005). Evolution of carbon sinks in a changing climate. Proceedings of the National Academy of Sciences, 102(32). https://doi.org/10.1073/pnas.0504949102

Fung, I., Prentice, K., & Matthews, E. (1983). Three-dimensional tracer model study of atmospheric CO_2: Response to seasonal exchanges with the terrestrial biosphere. Journal of Geophysical Research, 88(C2), 1281–1294. https://doi.org/10.1029/JC088iC02p01281

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., et al. (2019). Global atmospheric CO_2 inverse models converging on neutral tropical land exchange but diverging on fossil fuel and atmospheric growth rate. *Biogeosciences Discussions*, 1–25. https://doi.org/10.5194/bg-2018-384

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., et al. (2011). The Community Climate System Model Version 4. Journal of Climate, 24(19), 4973–4991. https://doi.org/10.1175/2011JCLI4083.1

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., et al. (2013). Enhanced Seasonal Exchange of CO_2 by Northern Ecosystems Since 1960. Science, 341(6150). https://doi.org/10.1126/science.1239207

Guo, C., Bentsen, M., Bethke, I., Ilicak, M., Tjiputra, J., Toniazzo, T., et al. (2019). Description and evaluation of NorESM1-F: a fast version of the Norwegian Earth System Model (NorESM). Geoscientific Model Development, 12(1), 343–362. https://doi.org/10.5194/gmd-12-343-2019

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., et al. (2002). Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. Nature. https://doi.org/10.1038/415626a

Hu, Q., Li, T., Deng, X., Wu, T., Zhai, P., Huang, D., Fan, X., Zhu, Y., Lin, Y., et al. (2022). Intercomparison of global terrestrial carbon fluxes estimated by MODIS and Earth system models. *Science of The Total Environment*, *810*, 152231. https://doi.org/10.1016/j.scitotenv.2021.152231

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Oda, T., Tans, P., et al. (2020). CarbonTracker CT2019B. NOAA Global Monitoring Laboratory. https://doi.org/10.25925/20201008

Jin, Y., Keeling, R. F., Morgan, E. J., Ray, E., Parazoo, N. C., and Stephens, B. B. (2021). A mass-weighted isentropic coordinate for mapping chemical tracers and computing atmospheric inventories. *Atmospheric Chemistry and Physics*, *21(1)*, 217–238. https://doi.org/10.5194/acp-21-217-202

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., et al. (2019). The FLUXCOM ensemble of global land-atmosphere energy fluxes. [Dataset] Scientific Data, 6(1). https://doi.org/10.1038/s41597-019-0076-8

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020

Keeling, C. D., Chin, J. F. S., & Whorf, T. P. (1996). Increased activity of northern vegetation inferred from atmospheric CO_2 measurements. Nature, 382(6587). https://doi.org/10.1038/382146a0

Keeling, C. D., Adams, J. A., Ekdahl, C. A., & Guenther, P. R. (1976). Atmospheric carbon dioxide variations at the South Pole. Tellus, 28(6), 552–564. https://doi.org/10.1111/j.2153-3490.1976.tb00702.x

Keeling, R. F., & Manning, A. C. (2014). Studies of Recent Changes in Atmospheric O2 Content. In Treatise on Geochemistry. Elsevier. https://doi.org/10.1016/B978-0-08-095975-7.00420-4

Keeling, R. F., & Graven, H. D. (2021). Insights from Time Series of Atmospheric Carbon Dioxide and Related Tracers. *Annual Review of Environment and Resources*, *46*(1), 85–110. https://doi.org/10.1146/annurev-environ-012220-125406

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., et al. (2020). GISS-E2.1: Configurations and Climatology. Journal of Advances in Modeling Earth Systems, 12(8). https://doi.org/10.1029/2019MS002025

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., et al. (2013). Atmospheric carbon dioxide variability in the community earth system model: Evaluation and transient dynamics during the twentieth and twenty-first centuries. Journal of Climate. https://doi.org/10.1175/JCLI-D-12-00589.1

Khatiwala, S., Primeau, F., & Hall, T. (2009). Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean. Nature, 462(7271). https://doi.org/10.1038/nature08526

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., 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

Kumar, J., Hoffman, F., Hargrove, W., & Collier, N. (2016). Understanding the representativeness of FLUXNET for upscaling carbon flux from eddy covariance measurements. Earth System Science Data Discussions, 1–25. https://doi.org/10.5194/essd-2016-36

Lee, W.-L., Wang, Y.-C., Shiu, C.-J., Tsai, I., Tu, C.-Y., Lan, Y.-Y., et al. (2020). Taiwan Earth System Model Version 1: description and evaluation of mean state. Geoscientific Model Development, 13(9), 3887–3904. https://doi.org/10.5194/gmd-13-3887-2020

Lee, Wei-Liang; Liang, Hsin-Chien (2020). AS-RCEC TaiESM1.0 model output prepared for CMIP6 CMIP historical.*Version 20200616* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.9755 .

Liptak, J., Keppel-Aleks, G., & Lindsay, K. (2017). Drivers of multi-century trends in the atmospheric CO&lt;sub&gt;2&lt;/sub&gt; mean annual cycle in a prognostic ESM. Biogeosciences, 14(6), 1383–1401. https://doi.org/10.5194/bg-14-1383-2017

Long, M. C., Lindsay, K., Peacock, S., Moore, J. K., & Doney, S. C. (2013). Twentieth-Century Oceanic Carbon Uptake and Storage in CESM1(BGC)*. Journal of Climate, 26(18), 6775–6800. https://doi.org/10.1175/JCLI-D-12-00184.1

Lovato, Tomas; Peano, Daniele (2020). CMCC CMCC-CM2-SR5 model output prepared for CMIP6 CMIP historical.*Version 20200527* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.3825 .

Lovato, Tomas; Peano, Daniele; Butenschön, Momme (2021). CMCC CMCC-ESM2 model output prepared for CMIP6 CMIP historical. *Version 20201225* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.13195 .

Mao, J., Thornton, P. E., Shi, X., Zhao, M., & Post, W. M. (2012). Remote sensing evaluation of CLM4 GPP for the period 2000-09. Journal of Climate, 25(15), 5327–5342. https://doi.org/10.1175/JCLI-D-11-00401.1

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its

Response to Increasing CO_2. Journal of Advances in Modeling Earth Systems, 11(4), 998–1038. https://doi.org/10.1029/2018MS001400

NASA Goddard Institute for Space Studies (NASA/GISS) (2018). NASA-GISS GISS-E2.1G model output prepared for CMIP6 CMIP historical. *Version 20181015* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.7127 .

NASA Goddard Institute for Space Studies (NASA/GISS) (2019). NASA-GISS GISS-E2.1H model output prepared for CMIP6 CMIP historical. *Version 20190328* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.7128 .

Neubauer, David; Ferrachat, Sylvaine; Siegenthaler-Le Drian, Colombe; Stoll, Jens; et al. (2019). HAMMOZ-Consortium MPI-ESM1.2-HAM model output prepared for CMIP6 CMIP historical. *Version 20191124* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.5016 .

Patra, P. K., Takigawa, M., Watanabe, S., Chandra, N., Ishijima, K., & Yamashita, Y. (2018). Improved chemical tracer simulation by MIROC4.0-based atmospheric chemistry-transport model (MIROC4-ACTM). Scientific Online Letters on the Atmosphere. https://doi.org/10.2151/SOLA.2018-016

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. https://doi.org/10.1073/pnas.0708986104

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

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y., & Field, C. B. (1997). The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. Global Biogeochemical Cycles, 11(4), 535–560. https://doi.org/10.1029/97GB02268

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., et al. (2004). The oceanic sink for anthropogenic CO_2. Science, 305(5682), 367–371. https://doi.org/10.1126/science.1097403

Santoni, G. W., Daube, B. C., Kort, E. A., Jiménez, R., Park, S., Pittman, J. V, et al. (2014). Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite-CO_2 , CH_4 , N_2O, and CO-during the CalNex and HIPPO campaigns. Atmos. Meas. Tech, 7, 1509–1526. https://doi.org/10.5194/amt-7-1509-2014

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., et al. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. Nature, 414(6860). https://doi.org/10.1038/35102500

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., et al. (2019). Quantifying the Impact of Atmospheric Transport Uncertainty on CO_2 Surface Flux Estimates. Global Biogeochemical Cycles, 33(4), 484–500. https://doi.org/10.1029/2018GB006086

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., et al. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. Geoscientific Model Development, 13(12), 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Seland, Øyvind; Bentsen, Mats; Oliviè, Dirk Jan Leo; Toniazzo, Thomas; et al. (2019). NCC NorESM2-LM model output prepared for CMIP6 CMIP historical. *Version20190917* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.8036 .

Simpson, I. R., McKinnon, K. A., Davenport, F. v., Tingley, M., Lehner, F., al Fahad, A., & Chen, D. (2021). Emergent constraints on the large scale atmospheric circulation and regional hydroclimate: do they still work in CMIP6 and how much can they actually constrain the future? *Journal of Climate*, 1–62. https://doi.org/10.1175/JCLI-D-21-0055.1

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., et al. (2007). Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2. Science. https://doi.org/10.1126/science.1137004

Stephens, B. B., Morgan, E. J., Bent, J. D., Keeling, R. F., Watt, A. S., Shertz, S. R., & Daube, B. C. (2021). Airborne measurements of oxygen concentration from the surface to the lower stratosphere and pole to pole. Atmospheric Measurement Techniques, 14(3), 2543–2574. https://doi.org/10.5194/AMT-14-2543-2021

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). Geoscientific Model Development, 12(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Swart, Neil Cameron; Cole, Jason N.S.; Kharin, Viatcheslav V.; Lazare, Mike; et al. (2019). CCCma CanESM5 model output prepared for CMIP6 CMIP historical. *Version 20190501* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.3610 .

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., et al. (2015). Seasonal climatology of CO_2 across North America from aircraft measurements in the NOAA/ESRL Global Greenhouse Gas Reference Network. Journal of Geophysical Research: Atmospheres, 120(10). https://doi.org/10.1002/2014JD022591

Tans, P. P., Fung, I. Y., & Takahashi, T. (1990). Observational constraints on the global atmospheric CO_2 budget. Science. https://doi.org/10.1126/science.247.4949.1431

Thompson, R. L., Patra, P. K., Chevallier, F., Maksyutov, S., Law, R. M., Ziehn, T., et al. (2016). Top–down assessment of the Asian carbon budget since the mid 1990s. *Nature Communications*, *7*(1), 10724. https://doi.org/10.1038/ncomms10724

Thompson, C. R., Wofsy, S. C., Prather, M. J., Newman, P. A., Hanisco, T. F., et al.(2021). THE NASA ATMOSPHERIC TOMOGRAPHY (ATom) MISSION: Imaging the Chemistry of the Global Atmosphere. *Bulletin of the American Meteorological Society*, 1–53. https://doi.org/10.1175/BAMS-D-20-0315.1

Thompson, C. R., Wofsy, S. C., Prather, M. J., Newman, P. A., Hanisco, T. F., Ryerson, T. B., Fahey, D. W., Apel, E. C., Brock, C. A., Brune, W. H., Froyd, K., Katich, J. M., Nicely, J. M., Peischl, J., Ray, E., Veres, P. R., Wang, S., Allen, H. M., Asher, E., … Zeng, L. (2022). The NASA Atmospheric Tomography (ATom) Mission: Imaging the Chemistry of the Global Atmosphere. *Bulletin of the American Meteorological Society*, *103*(3), E761–E790. https://doi.org/10.1175/BAMS-D-20-0315.1

K.W. Thoning, A.M. Crotwell, and J.W. Mund (2022), Atmospheric Carbon Dioxide Dry Air Mole Fractions from continuous measurements at Mauna Loa, Hawaii, Barrow, Alaska, American Samoa and South Pole. 1973-2021, Version 2022-05 National Oceanic and

Atmospheric Administration (NOAA), Global Monitoring Laboratory (GML), Boulder, Colorado, USA https://doi.org/10.15138/yaf1-bk21

Tjiputra, J. F., Roelandt, C., Bentsen, M., Lawrence, D. M., Lorentzen, T., Schwinger, J., et al. (2013). Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM). Geoscientific Model Development, 6(2), 301–325. https://doi.org/10.5194/gmd-6-301-2013

Valentini, R., Matteucci, G., Dolman, A. J., Schulze, E. D., Rebmann, C., Moors, E. J., et al. (2000). Respiration as the main determinant of carbon balance in European forests. Nature, 404(6780), 861–865. https://doi.org/10.1038/35009084

Verma, S., Marshall, J., Gerbig, C., Rödenbeck, C., & Uwe Totsche, K. (2017). The constraint of CO_2 measurements made onboard passenger aircraft on surface-atmosphere fluxes: The impact of transport model errors in vertical mixing. Atmospheric Chemistry and Physics, 17(9), 5665–5675. https://doi.org/10.5194/acp-17-5665-2017

Volodin, E. M., Dianskii, N. A., & Gusev, A. V. (2010). Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. Izvestiya, Atmospheric and Oceanic Physics, 46(4), 414–431. https://doi.org/10.1134/S000143381004002X

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., et al. (2011). MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. Geoscientific Model Development, 4(4), 845–872. https://doi.org/10.5194/gmd-4-845-2011

Welp, L. R., Keeling, R. F., Meijer, H. A. J., Bollenbacher, A. F., Piper, S. C., Yoshimura, K., et al. (2011). Interannual variability in the oxygen isotopes of atmospheric CO_2 driven by El Niño. Nature, 477(7366), 579–582. https://doi.org/10.1038/nature10421

Williamson, D. B., & Sansom, P. G. (2019). How Are Emergent Constraints Quantifying Uncertainty and What Do They Leave Behind? *Bulletin of the American Meteorological Society*, *100*(12), 2571–2588. https://doi.org/10.1175/BAMS-D-19-0131.1

Williamson, M. S., Thackeray, C. W., Cox, P. M., Hall, A., Huntingford, C., & Nijsse, F. J. M. M. (2021). Emergent constraints on climate sensitivities. *Reviews of Modern Physics*, *93*(2), 025004. https://doi.org/10.1103/RevModPhys.93.025004

Wieners, Karl-Hermann; Giorgetta, Marco; Jungclaus, Johann; Reick, Christian; et al. (2019). MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP historical. *Version 20190929* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.6595 .

Wofsy, S. C., The HIPPO Science Team and Cooperating Modellers and Satellite Teams, (2011). HIAPER Pole-to-Pole Observations (HIPPO): Fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. https://doi.org/10.1098/rsta.2010.0313

Wofsy, S. C., Daube, B. C., Jimenez, R., Kort, E., Pittman, J. V., Park, S., et al. (2017). HIPPO Merged 10-second Meteorology, Atmospheric Chemistry, and Aerosol Data. Version 1.0. [Dataset] UCAR/NCAR - Earth Observing Laboratory. Retrieved from http://doi.org/10.3334/CDIAC/HIPPO_010

Wofsy, S.C., S. Afshar, H.M. Allen, E.C. Apel, E.C. Asher, B. Barletta, J. Bent, et al.(2021). ATom: Merged Atmospheric Chemistry, Trace Gases, and Aerosols, Version 2. [Dataset] ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1925

Yang, Z., Washenfelder, R. A., Keppel-Aleks, G., Krakauer, N. Y., Randerson, J. T., Tans, P. P., et al. (2007). New constraints on Northern Hemisphere growing season net flux. Geophysical Research Letters. https://doi.org/10.1029/2007GL029742

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., et al. (2020). The Australian Earth System Model: ACCESS-ESM1.5. Journal of Southern Hemisphere Earth Systems Science, 70(1), 193. https://doi.org/10.1071/ES19035

Ziehn, Tilo; Chamberlain, Matthew; Lenton, Andrew; Law, Rachel; Bodman, Roger; et al. (2019). CSIRO ACCESS-ESM1.5 model output prepared for CMIP6 CMIP historical. *Version 20191115* Earth System Grid Federation. doi:https://doi.org/10.22033/ESGF/CMIP6.4272

2.7 Supporting Information

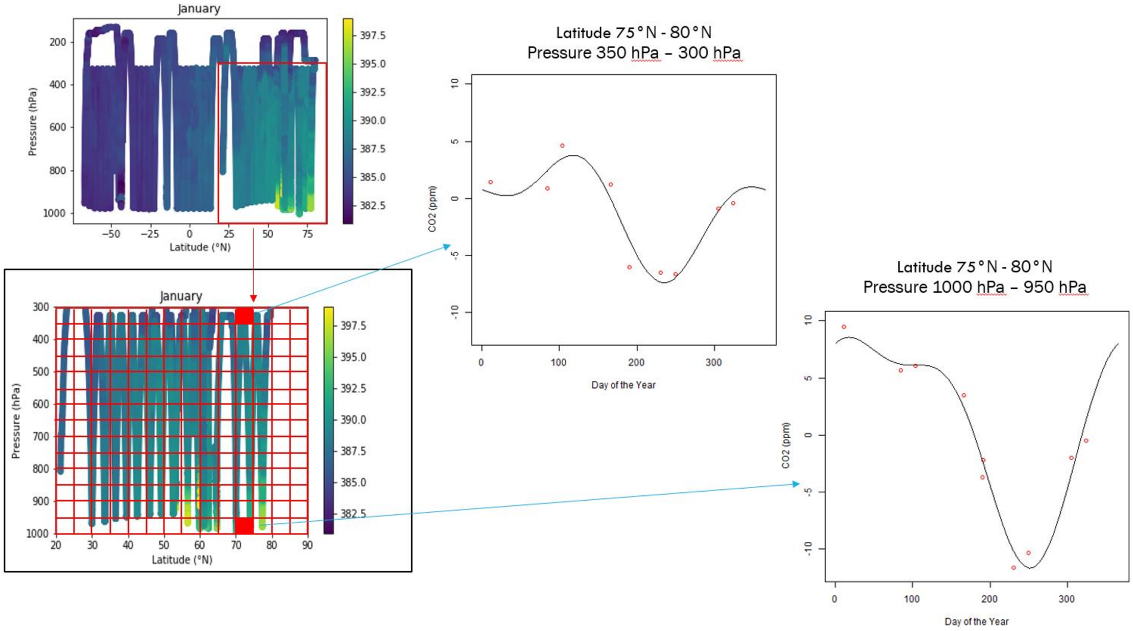


Figure 2.8 (S1) Diagram of bin fitting process. Atmospheric CO_2 observations from the HIPPO and ATom campaigns are averaged by day, then we take only measurements north of 20° and with pressures greater than 300 hPa. This data is binned by 5° in latitude and 50 hPa in pressure. A second-order harmonic is fit to all data within a given bin. The seasonal cycle of a bin depends on both latitude and pressure.

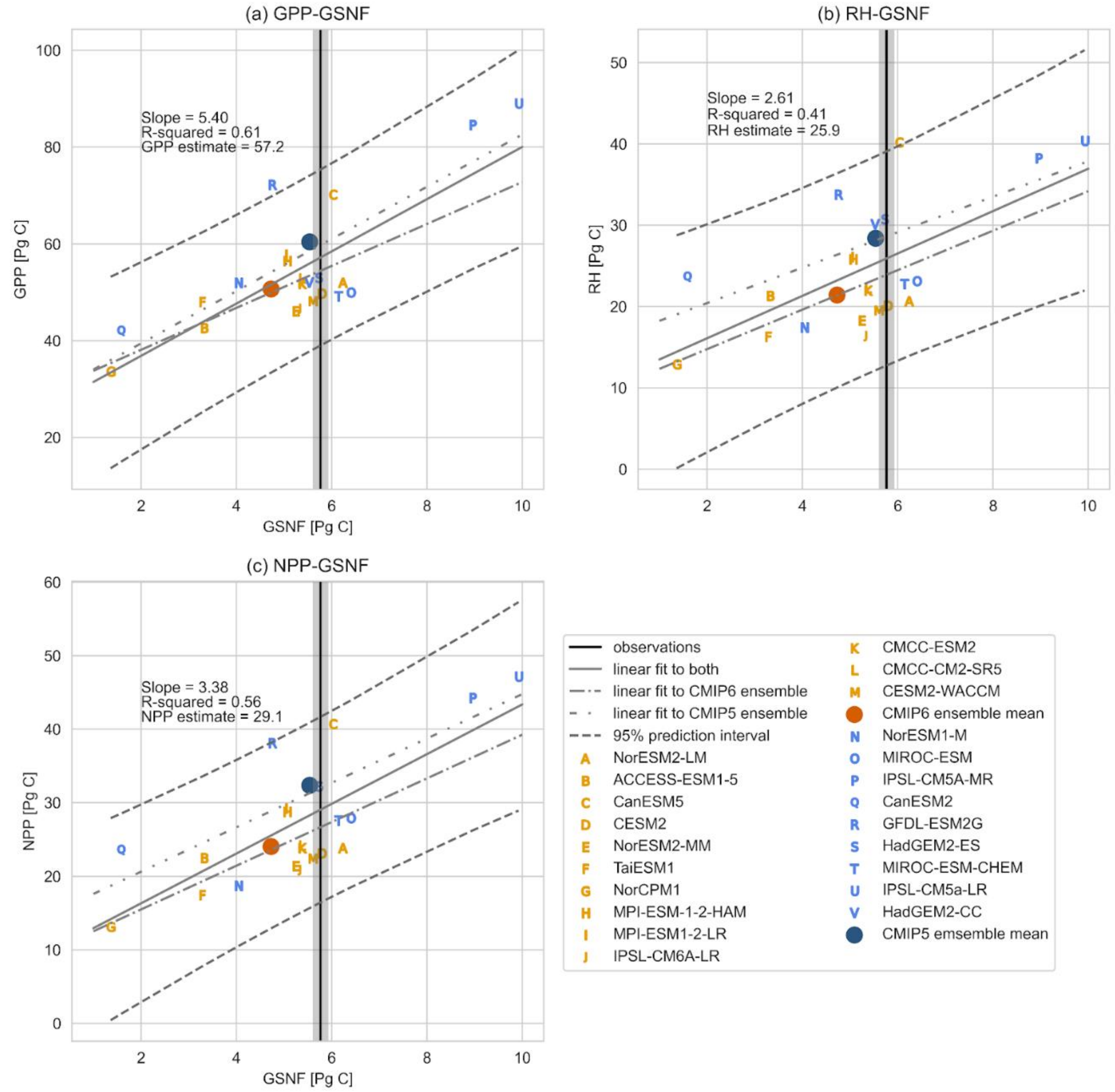


Figure 2.9 (S2) Calculating the GSNF from each CMIP model by prescribing the start and end day from the observations, we are able to evaluate how well the models match the observations. The primary difference in the results using this approach is a slight reduction in the magnitude of the GSNF for most of the CMIP5 and CMIP6 models, a less strong correlation between GSNF and GPP, RH, and NPP, and a slight increase in the inferred values of northern GPP, RH, and NPP.

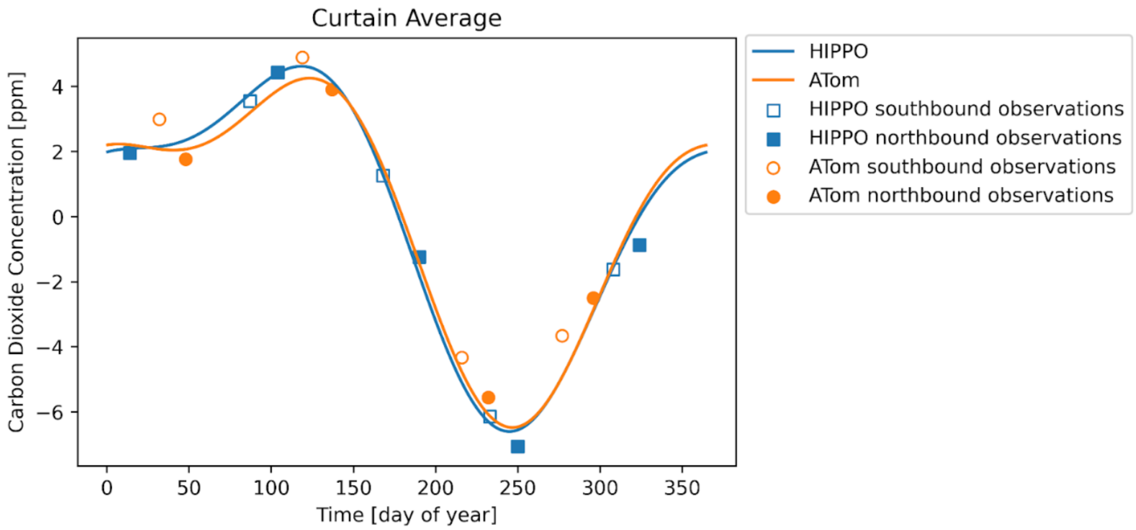


Figure 2.10 (S3) HIPPO and ATom curtain averages. Two-harmonic fit to detrended average carbon dioxide concentration in ppm as a function of day of year for HIPPO (blue) and ATom (orange) flight campaigns in the atmospheric curtain between 20°N and 90°N in latitude and between 1000 hPa and 300 hPa in pressure. Blue squares are each HIPPO campaign, with the southbound leg being empty and the northbound leg being filled. Orange circles each ATom campaign, with the southbound leg being empty and the northbound leg being filled. The blue squares and orange circles are found by filtering, detrending, interpolating and extrapolating to get a full altitude and latitude slice, then taking a pressure and cosine of latitude weighted average. The lines are the average of all fits to individual latitude-pressure bins with the annual mean removed for each campaign individually.

Chapter 3 **Representativeness and Variation of Seasonal Total Column CO_2**

3.1 Introduction

Understanding likely feedbacks between climate and the carbon cycle requires accurate inferences of net carbon fluxes and the processes that control them over a range of spatial scales. At the broadest spatial scales, we must be able to accurately partition the fate of anthropogenic carbon among atmosphere, ocean, and land. Yet in the most recent accounting of the global carbon budget, there was an imbalance between estimated total emissions and estimated changes in carbon stored in the atmosphere, ocean, and land of -0.8 Gt C yr^{-1} for 2020 (Friedlingstein et al., 2022). At finer spatial scales, knowledge of surface-atmosphere carbon exchange may provide insight about how ecosystems respond to heterogeneous patterns of temperature, precipitation, and other environmental factors.

While many methods are used to estimate land-atmosphere carbon fluxes at various spatial scales, there are spatiotemporal gaps in our knowledge of carbon fluxes. At hemispheric scales, atmospheric measurements of carbon dioxide (CO_2) can constrain biospheric carbon exchange. For example, in chapter 2, aircraft measurements of atmospheric CO_2 were used to infer the growing season net flux (GSNF) for the northern extratropics. However, this approach is limited to large spatial scales and to periods in time when aircraft campaigns are flown.

Atmospheric CO_2 observations can be used to infer fluxes at regional to continental spatial scales using atmospheric inversion or data assimilation techniques. In inversion systems, carbon fluxes are optimized so that the resulting CO_2 fields, produced by running those fluxes through an atmospheric transport model, match a distribution of measured values of atmospheric

CO_2. However, the estimates of carbon exchange found using inversion systems tend to be highly uncertain at sub-continental spatial scales. An experiment (TransCom 3) undertaken to explore inversion estimates of carbon fluxes showed that the spread between model estimates was larger than the average of their estimated flux uncertainty in the northern extratropics, with the spread being largest during the growing season (Gurney et al., 2004). Tolk et al., (2011) used four different inversion methods to estimate the spatiotemporal distribution of NEE in the Netherlands and found that all four methods struggled to reproduce the true distribution. Another study by Xiao et al., (2014) sought to quantify the uncertainty of mean NEE estimates for 2001-2007 in the upper Midwestern United States and found an uncertainty of over 25%. The variation in estimated fluxes in the northern hemisphere is primarily due to the response to background biosphere exchange (Gurney et al., 2003).

While inferences about fluxes from atmospheric CO_2 tend to stall at the regional spatial scale, direct observations of net carbon fluxes can be made on the scale of a few kilometers using eddy covariance flux towers. Eddy covariance provide continuous measurements of net ecosystem exchange (NEE) (Baldocchi 2001), but scaling these observations up can be problematic because of the sparseness of the network and small spatial footprint of each tower (Baldocchi et al., 2003). Efforts have been made to scale up the flux tower measurements to global-scale estimates of NEE using machine learning (Jung et al., 2019). However, current estimates of NEE using machine learning do not accurately capture regional-scale carbon sinks. For example, Jung et al. (2020) showed that FLUXCOM overestimated mean annual NEE in the tropics by hundreds of grams of carbon per square meter compared to inversion-based and dynamic global vegetation model-based estimates and in Chapter 2 we showed that NH annual

GPP was on the high end of FLUXCOM estimates. Thus, the biases in FLUXCOM call into question its trustworthiness at grid scale to regional-scale carbon exchange.

Remote sensing observations may provide insights that allow us to scale from hemispheric to grid scale level. The Orbiting Carbon Observatory 2 (OCO-2) retrieves XCO_2 at dense spatial coverage but is sparsely sampled in time with retrievals at any location only obtained once every 16 days or so. OCO-2 also measures solar induced chlorophyll fluorescence (SIF) which has been used as a proxy for ecosystem productivity and may shed light on site-to-site variation in seasonal carbon exchange.

Comparing atmospheric measurements of CO_2 taken at different frequencies and locations can help increase understanding of local and regional carbon seasonal cycles. The Total Carbon Column Observing Network (TCCON) of ground-based Fourier transform spectrometers retrieves column-averaged dry-air mole fractions of CO_2 (XCO_2). The TCCON has sparse spatial coverage with only about 30 locations worldwide but is densely sampled in time with retrievals obtained every few minutes. The dense temporal sampling of TCCON observatories allows for analysis of the diurnal carbon cycle, which may shed light on local carbon exchange.

In addition to XCO_2 measurements, vertical profiles of atmospheric CO_2 have been obtained during the HIPPO and ATom aircraft campaigns. These profiles are sampled at each latitude over the remote oceans within each season. Comparison of these three types of observations can shed light on how differences in spatial and temporal sampling of atmospheric CO_2 affect the northern extratropical (NET) seasonal cycle.

In this study, we explore whether remote sensing observations of XCO2 can provide hemispheric constraints consistent with that from HIPPO and ATOM, and whether we can account for site-level or grid scale-level patterns of seasonal carbon exchange using ancillary data such as the diurnal cycle or productivity from SIF. Our ultimate goal is to work toward the ability to have robust and self-consistent constraints on carbon exchange across scales -- from hemispheric to grid scale. We describe the data and methods used to address these questions in section 3.2. Section 3.3.1 will present the results related to the hemispheric seasonal cycle and section 3.3.2 will present the results related to local and regional seasonal cycles. Section 3.4 will discuss the broader implications of the results presented in sections 3.3.

3.2 Data and Methods

3.2.1 TCCON

The TCCON is a network of ground-based Fourier transform spectrometers. The network provides multi-year, nearly continuous sunlit observations at about 30 sites located around the globe (Wunch et al., 2011). Sites within the network retrieve total column average, dry air mole fractions (DMFs) of numerous gases including CO_2. Total column abundances are retrieved from direct solar absorption spectra taken by ground-based Bruker 125HR (high resolution) Fourier transform spectrometers (FTS) using optimal estimation via the GGG2020 algorithm (TCCON Team, 2022; Laughner et al., 2022). The column-average DMF is then calculated by dividing the total column abundance of the gas of interest with that of atmospheric oxygen (O_2), whose mixing ratio is well known (0.2095) and whose column abundance correlates with surface pressure (Wunch et al. 2010). Our study focuses on CO_2, which is retrieved every few minutes in the near infrared (Wunch et al., 2011) during daylight hours. The reported TCCON column average DMF (XCO_2) data are calibrated to the World Meteorological Organization (WMO) standard with absolute uncertainty of ~0.3% (Washenfelder et al., 2006; Wunch et al., 2010) using aircraft profile data.

We use XCO_2 retrievals from 15 sites located in the northern hemisphere, spanning 18.5 °N to 67.37 °N (Table 1). The sites we used are all part of the TCCON GGG2020 dataset and the data used in this study was downloaded from the TCCON archive at https://tccondata.org/.

In addition to the XCO_2 retrievals from TCCON, we use the TCCON prior vertical profiles. These profiles are calculated using the GGG2020 algorithm. The reported prior profile is such that the RMS fitting residuals between the observed spectra and the modeled spectra are minimized. We focus on the stratospheric portion of the prior.

3.2.2 OCO-2

NASA's first dedicated CO_2 observing satellite, OCO-2, was launched in 2014. The satellite flies in a sun-synchronous polar orbit at an altitude of 705 km. OCO-2 has an overpass of roughly 13:30 local time and repeats on a 16-day cycle (Eldering et al., 2017). Measurements from OCO-2 are made using three grating spectrometers which measure reflected sunlight in the near infrared at 0.765, 1.61, and 2.06 µm. OCO-2 has a swatch width of 10.4 km and obtains 8 across track footprints, resulting in a sampling footprint of 1.29 x 2.25 square kilometer. OCO-2 observes in three scanning modes, nadir, glint, and target, but we only use data obtained while in nadir and glint modes. The nadir mode is used over land as it provides slightly higher spatial resolution and better signal-to-noise ratio over land. The glint mode maximizes reflected light over water surfaces by pointing the instrument at the glint angle and is used over both land and water. The target mode allows for validation of data with a ground-based sites by adjusting the instrument angle to point toward a specific location, such as the TCCON sites described above.

Table 3.1 Location, observational periods, and data references for TCCON sites analyzed in this study.

TCCON site	Site id	Location	Observational Period	Data Reference
Bremen, Germany	br	53.10°N, 8.85°E	2009/01/06-2021/06/24	Notholt et al. 2019
Burgos, Ilocos Norte, Philippines	bu	18.53°N, 120.65°E	2017/03/03-2021/04/30	Morino et al. 2017a
California Institute of Technology, Pasadena, California, USA	ci	34.14°N, 118.13°W	2012/09/20-2022/03/01	Wennberg et al. 2017a
East Trout Lake, Canada	et	54.35°N, 140.99°W	2016/10/03-2022/05/28	Wunch et al. 2020
Garmisch, Germany	gm	47.48°N, 11.06°E	2007/07/16-2021/10/18	Sussmann et al. 2017
Izana, Tenerife, Spain	iz	28.31°N, 16.49°W	2014/01/02-2022/04/27	Blumenstock et al. 2017
Saga, Japan	js	33.24°N, 130.29°E	2011/07/28-2021/06/30	Shiomi et al. 2022
Karlsruhe, Germany	ka	49.10°N, 8.44°E	2014/01/15-2021/12/22	Hase et al. 2017
Lamont, Oklahoma, USA	oc	36.60°N, 97.49°W	2008/07/06-2022/05/01	Wennberg et al. 2017b
Orleans, France	or	47.97°N, 2.11°E	2009/08/29-2021/03/03	Warneke et al. 2017
Park Falls, Wisconsin, USA	pa	45.94°N, 90.27°W	2004/05/26-2022/02/28	Wennberg et al. 2017c
Paris, France	pr	48.85°N, 2.36°E	2014/09/23-2021/06/16	Te et al. 2017
Rikubetsu, Hokkaido, Japan	rj	43.46°N, 143.77°E	2014/06/24-2021/06/30	Morino et al. 2017b
Sodankyla, Finland	so	67.37°N, 26.63°E	2018/03/05-2021/10/18	Kivi et al. 2017
Tsukuba, Ibaraki, Japan	tk	36.05°N, 140.12°E	2014/03/28-2021/06/28	Morino et al. 2017c

The OCO-2 XCO_2 product is produced using a Bayesian optimal estimation retrieval algorithm (Atmospheric CO_2 Observations from Space (ACOS), O'Dell et al., 2018) which estimates the vertical column of both CO_2 and O_2 while adjusting other elements of the retrieval state vector, similar to the approach used in the GGG2020 algorithm. Soundings are bias corrected based on their relationship with TCCON validation data (Wunch et al., 2017), resulting in errors that are generally less than 1 ppm for retrieved XCO_2 (Eldering et al., 2017). In this study, we use the bias corrected XCO_2 values reported in the version 10 Level 2 Lite files.

In this study, we also analyzed the Solar-Induced Chlorophyll Florescence (SIF) reported by OCO-2. SIF is a proxy for gross primary productivity. The OCO-2 SIF product is found by in-filling of solar Fraunhofer lines around 757 nm and 771 nm (Sun et al., 2018). We average the SIF values at 757 nm and 771 nm multiplied by a factor of 1.5 because the value at 757 nm is typically ~1.5 times larger than that at 771 nm.

3.2.3 Seasonal cycle and SCA calculation

The TCCON and OCO-2 XCO_2 retrievals are used to calculate the seasonal cycle of XCO_2. We first filter the data to ensure high quality soundings are used by removing measurements taken at each TCCON site when the solar zenith angle is larger than 75 degrees to minimize spectroscopic errors (Fig. 3.2). For OCO-2 the data is then aggregated to a 5° by 5° global grid, and grid cells where data is collected for fewer than 6 months out of the year are removed to ensure there are enough soundings in each grid cell to fully constrain the seasonal cycle in that cell.

We additionally calculate seasonal cycles and seasonal cycle amplitudes using the most recent Carbon Tracker release (CT2019B, Jacobson et al., 2020), described in section 2.2.3. The CT2019B output is not filtered, but it is aggregated to a 5° by 5° global grid, similarly to OCO-2.

This global grid is then sampled in multiple ways. We sample CT2019B at the locations of TCCON sites and at the locations HIPPO and ATom measurements were taken over the period 2009-2018. These are referred to as CT2019B sampled like TCCON (CT like TCCON) and CT2019B sampled like HIPPO and ATom continuously (CT like HIPTOM cont.). In addition to these, we also use the full CT2019B dataset between 25 and 55 degrees north, referred to as CT2019B (CT) and the CT2019B obspack output for HIPPO and ATom, referred to as CT2019B sampled like HIPPO and ATom (CT like HIPTOM).

For both TCCON and OCO-2 retrievals, as well as for CT2019B, the next step in the process of finding the seasonal cycle is to average the data to monthly means, resulting in a time series at monthly resolution. The data are then detrended by removing the long-term trend found by applying seasonal and trend decomposition using Loess (STL) to the NOAA Mauna Loa record. STL decomposes a time series into trend, seasonal, and remainder components through application of the loess smoother (Cleveland et al., 1990). We use STL because it is generally robust to outliers and allows for decomposition of time series with missing values (Cleveland et al., 1990). The detrended time series is then binned by month of the year and the mean is taken for each monthly bin resulting in the mean seasonal cycle. A second-order harmonic is fit to the mean seasonal cycle, and he seasonal cycle amplitude is found by subtracting the minimum value from the maximum value of the harmonic fit of the seasonal cycle.

3.2.4 Stratospheric adjustment

To calculate a seasonal cycle that is comparable to the seasonal cycle obtained using HIPPO and ATom measurements, the TCCON column must be adjusted to account for the influence of stratospheric CO_2 on the column average. We estimate the stratospheric component of the total column using the TCCON prior profile. The prior profile is integrated according to equation 6 ($VC_{G,ak} = \int_0^{P_s} \frac{f_G^{dry}(p) \cdot a(p)}{g(z(p),\phi) \cdot m_{air}^{dry} \cdot \left[1 + f_{H_2O}^{dry}(p) \cdot (m_{H_2O}/m_{air}^{dry})\right]} dp$) from Wunch et al. (2010) but integrated between 300 hpa and 0 hpa in pressure, instead of from the surface to zero pressure, to estimate the stratospheric contribution. The partial vertical column (VC) of CO_2 is divided by the partial column of O_2 to get the stratospheric column-averaged dry-air mole fraction (Eq. 1), whose seasonal cycle is shown by the dotted blue line in figure 3.1.

$$stratXCO_2 = 0.2095 \frac{stratVC_{CO_2}}{stratVC_{O_2}} \qquad (1)$$

The stratospheric XCO_2 is removed from the total column XCO_2 to get a resulting adjusted XCO_2 according to equation 2,

$$tropXCO_2 = \frac{XCO_2 - \frac{P_{strat}}{P_s}(stratXCO_2)}{1 - \frac{P_{strat}}{P_s}} \qquad (2)$$

where XCO_2 is the total column value, P_{strat} is the pressure at the floor of the stratosphere, and P_s is surface pressure. The ratio of P_{strat} and P_s is representative of the percent of the atmosphere that is contained in the stratosphere. Note, we use 300 hPa for P_{strat} for the analysis but did vary the pressure floor as part of the sensitivity analysis, described in section 3.4.1. The resulting adjusted column is then detrended and the seasonal cycle is found as described above (dashed orange line in figure 3.1).

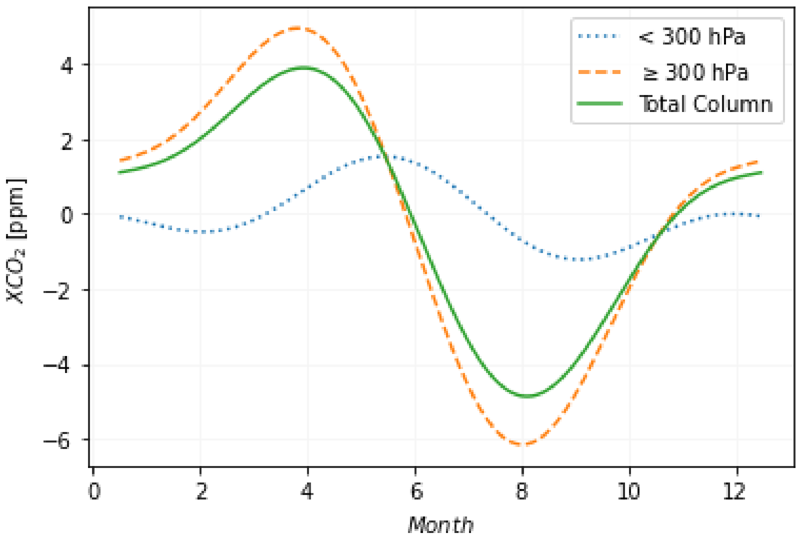


Figure 3.1 Second-order harmonic fit to the seasonal cycle of XCO2 for full column (solid green), stratosphere (dotted blue), and troposphere (dashed orange) at Park Falls, Wisconsin TCCON site.

The northern extratropical mean seasonal cycle is found by first removing data from sites south of 25 °N and north of 55 °N. These latitude bounds were chosen because all three data types used in the analysis are available within these bounds, allowing for direct comparison. The data are then aggregated into 5° latitude bins and the seasonal cycle is found within each bin as described above. We then take the area-weighted average of these zonal mean data, where weighting is by cosine of latitude. Again, a second-order harmonic is fit to the mean seasonal cycle and the seasonal cycle amplitude is found by subtracting the minimum value from the maximum value of the harmonic fit.

3.2.5 Diurnal cycle and DCA calculation

TCCON XCO_2 is temporally dense allowing for measurements of diurnal changes in atmospheric CO_2. When finding the diurnal cycle, we start by filtering anthropogenic fossil fuel emissions from urban TCCON sites (Paris, France; Saga, Japan; Pasadena, California, USA; and

Bremen, Germany) to remove the influence of local emissions on the diurnal cycle. Local pollution results in a diurnal covariation of CO_2 with CO, as documented by Wunch et al. (2009). This correlation is used to remove anthropogenic emissions as described in Appendix A of Keppel-Aleks et al. (2012), using the following equation:

$$CO_{2_{clean}} = CO_2 - \frac{CO - (\overline{CO} - 0.015\,ppm)}{EF}$$

where EF is the emission factor between CO and CO_2 found as the slope when CO_2 is regressed against CO. We found EF to be 0.04, resulting in a decrease of ~ 0.5 ppm for clean air relative to the unfiltered retrieval (S 1). The data are also filtered by solar zenith angle and the long-term trend is removed as described above in section 3.2.3.

The seasonal cycle obtained after filtering and detrending the data from a particular site is primarily due to terrestrial exchange (Tucker et al., 1986; Knorr and Heimann, 1995; Erickson et al., 1996). The data need to be deseasonalized to remove the known seasonal cycle, which is larger in amplitude and might otherwise obscure the smaller amplitude diurnal cycle. For each site, the entire time series is binned by day of the year and the mean value across all years is taken. The result is the mean seasonal cycle at daily resolution. This cycle is smoothed by taking a 7-day rolling average, and this rolling average is removed from the detrended XCO_2, leaving a residual that contains diurnal variation. The diurnal cycle (Fig. 3.2) is found by binning the residual by month of the year and hour of the day and taking the mean within each bin, removing any bin with less than 100 measurements. The diurnal cycle amplitude is then found by fitting a line to the slope of the diurnal cycle (Fig. 3.2). The slope of the line is multiplied by the average number of daylight hours for that month at that location, found at weatherspark.com, to give the diurnal cycle amplitude.

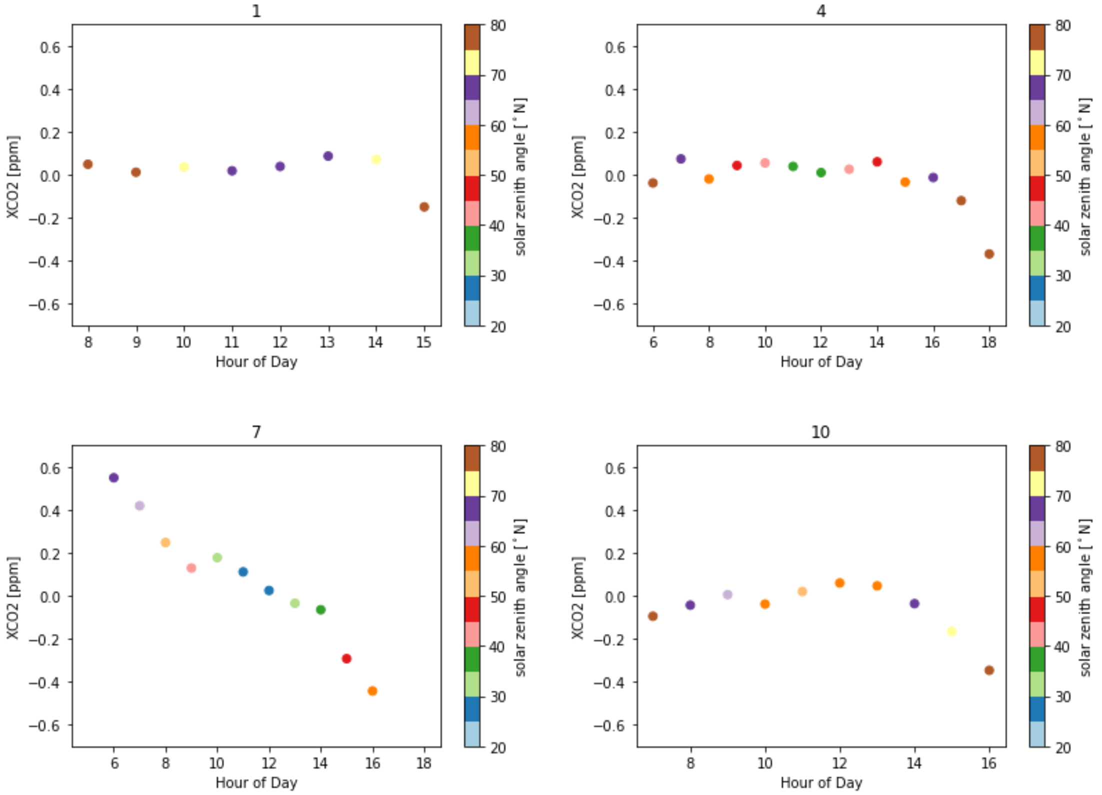


Figure 3.2 Plot of mean diurnal cycle in January, April, July, and October at Park Falls, Wisconsin. Spectroscopic errors can be seen at solar zenith angles larger than 75 degrees.

3.2.6 Equivalent Latitude

We also use an equivalent latitude of potential temperature at 700 hPa. To find potential temperatures we use the temperatures and pressures from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) reanalysis data. Similarly to OCO-2, the MERRA-2 data is aggregated to a 5° by 5° global grid. Within each grid cell, the monthly mean is taken, resulting in a time series at monthly resolution. The potential temperature at 700 hPa in each bin is found according to equation:

$$\theta = T\left(\frac{700}{P}\right)^{0.286}$$

The data is then binned by month and a mean taken in each bin to find the mean monthly potential temperature at 700 hPa.

3.3 Results

We compared curtain average seasonal cycles of atmospheric CO_2 calculated using column-average TCCON and OCO-2 measurements with those calculated using HIPPO and ATom aircraft measurements to determine the viability of using the TCCON to monitor the growing season net flux of carbon in the northern hemisphere. The results of these comparisons are reported in section 3.3.1 and discussed in section 3.4.1.

To better understand the site-to-site variation in the carbon seasonal cycle, we also analyzed the relationships between seasonal cycle amplitude, potential temperature, solar induced chlorophyll fluorescence, and diurnal cycle amplitude. The results of this analysis is reported in section 3.3.2 and discussed in section 3.4.2.

3.3.1 Hemispheric Seasonal Cycle

Figure 3.3 shows the curtain average seasonal cycles and SCAs between 25° N and 55° N for TCCON (both total column and adjusted column), OCO-2, and HIPPO/ATom. The total column TCCON seasonal cycle is similar in phase and amplitude to the OCO-2 seasonal cycle. There are only slight differences in the shape of the seasonal cycles, and the SCA of 6.9 ± 0.4 ppm for TCCON is only 0.1 ppm larger than the SCA of 6.8 ± 0.1 ppm for OCO-2. The TCCON seasonal cycle first crosses zero on day 173, just 3 days before the day 176 zero crossing for OCO-2.

In contrast, the adjusted-TCCON seasonal cycle does not show good agreement with the HIPPO/ATom seasonal cycle. The SCA of 8.3 ± 0.4 ppm for adjusted-TCCON is at least 0.7

ppm lower than the SCA of 10.0 ± 0.6 ppm for HIPPO/ATom. There is also significant disagreement in the day of first zero crossing with the first crossing occurring on day 171 for TCCON and on day 181 for HIPPO/ATom.

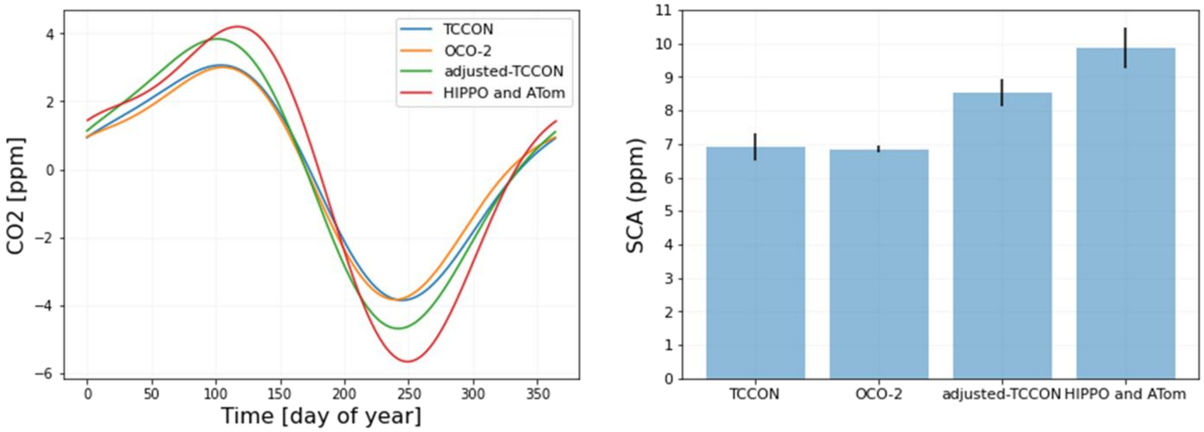


Figure 3.3 Curtain average seasonal cycles and SCA for 25-55 degrees north for TCCON, OCO-2, and HIPPO and ATom observations. Adjusted-TCCON has been adjusted to remove the stratospheric influence on the total column, leaving the tropospheric column.

To determine the reason for disagreement in the seasonal cycles for 25° N to 55° N calculated using the adjusted-TCCON observations and calculated using HIPPO and ATom measurements, we compare seasonal cycles calculated using atmospheric CO_2 simulations from CT2019B in figure 3.4. When CT2019B is sampled at the times and locations along the HIPPO and ATom flight tracks, the resulting seasonal cycle has a similar shape to the seasonal cycle resulting from taking the full zonal mean CO_2 from CT2019B but begins decreasing and reaches its minimum value slightly later in the year. The CT2019B sampled like HIPPO and ATom SCA of 9.6 ppm is 0.4 ppm larger than the 9.2 ppm SCA found when using the full zonal mean. This disagreement is decreased when CT2019B is sampled continuously between 2009 and 2018 along the HIPPO and ATom flight tracks. The shape, timing, and amplitude of the seasonal cycle for CT using the full zonal mean is nearly identical to the seasonal cycle for CT sampled like

HIPPO and ATom continuously, and there is only 0.2 ppm difference in their SCA with the continuously sampled CT2019B seasonal cycle having an amplitude of 9.4 ppm compared to the 9.2 ppm SCA of CT2019 using the full zonal mean.

In contrast, when CT is sampled like TCCON, the resulting seasonal cycle has similar shape and timing to the full zonal mean, but has a significantly smaller amplitude. Taking the tropospheric column at TCCON sites results in a SCA of 8.1 ppm, 1.1 ppm smaller than the SCA for the full zonal mean. We saw the same underestimate of the SCA from CT2019B sampled like TCCON when compared to the full zonal average when integrating the total column. The SCA of 5.9 ppm for CT2019B sampled like TCCON is 0.7 ppm smaller than the SCA of 6.6 ppm for CT2019B. These results indicate that the HIPPO and ATom seasonal cycle, while biased slightly high due to interannual variability, is reflective of the full zonal average. These results also indicate that TCCON is biased low and is not reflective of the full zonal average.

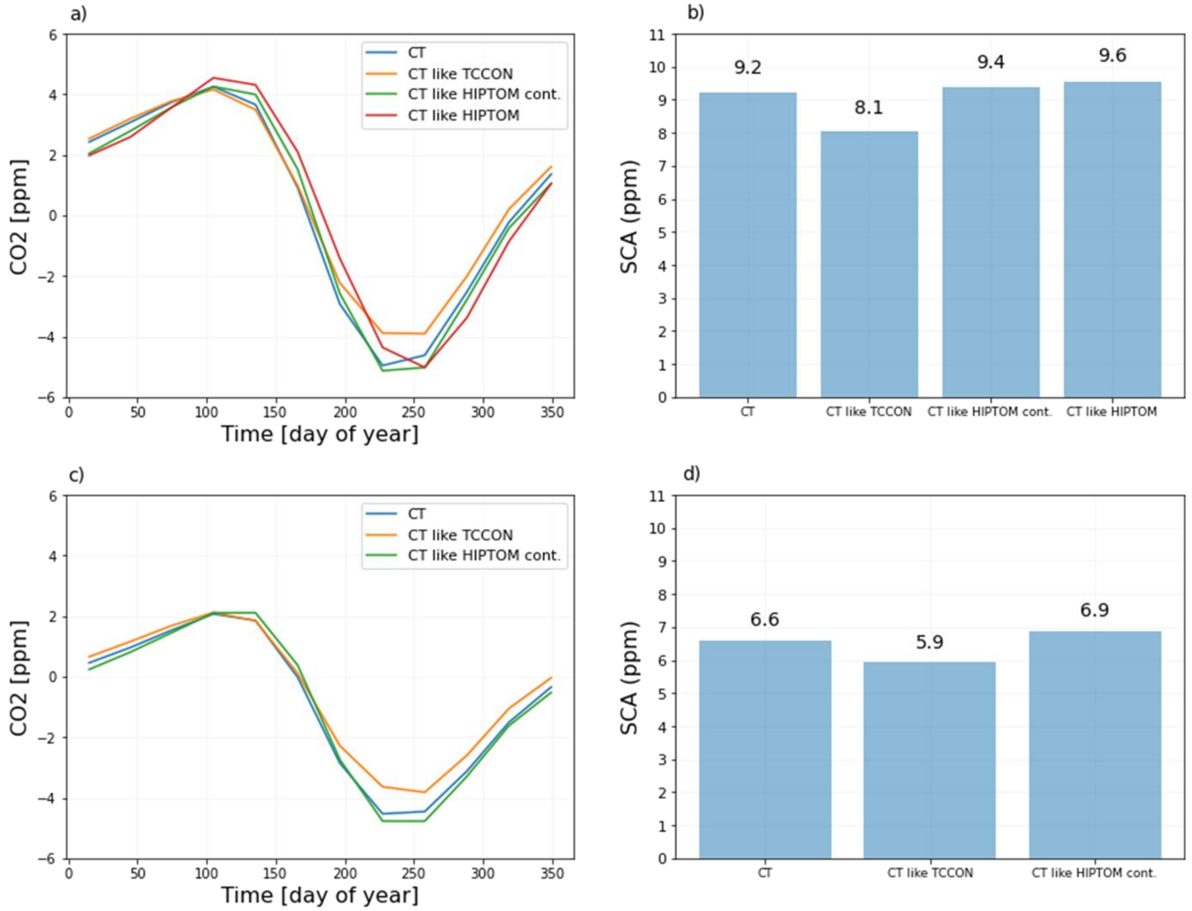


Figure 3.4 Curtain average seasonal cycles and SCA between 25 and 55 degrees north calculated using CT2019B. (a,b) Integration of the partial column below 300 hPa in atmospheric pressure and (c,d) integration of the total column.

3.3.2 Site-to-site Seasonal Cycles

The sparse spatial sampling of the TCCON might be leading to the difference seen when the seasonal cycle calculated using TCCON observations is compared to the seasonal cycle calculated using HIPPO and ATom measurements. To better understand the site-to-site variation in carbon seasonal cycles, we overlaid SCAs calculated from TCCON observations on the SCAs calculated from OCO-2 observations aggregated to a 5 degree by 5 degree grid (Fig. 3.5, a). In general, the TCCON-derived site SCA agrees well with the OCO-2-derived SCA within the region in which the TCCON site lies and we see a similar pattern of increasing SCA with latitude in both datasets. However, the TCCON-derived site SCA is always larger than the OCO-2-derived SCA, and 10 of the 14 TCCON sites analyzed had a SCA that is at least 1 ppm larger than the OCO-2 SCA in that grid cell (Fig. 3.5, b). The largest differences, which are as large as 2-3 ppm or around 30% of the SCA, are seen for the sites at Park Falls, USA; Paris, France; Orleans, France; and Rikubetsu, Japan.

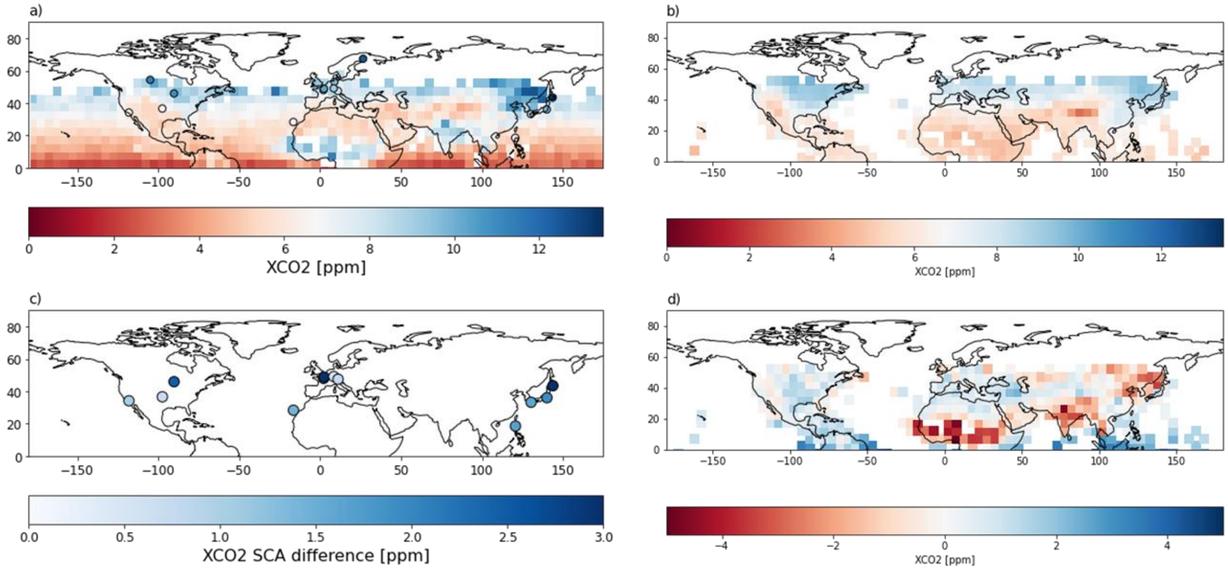


Figure 3.5 (a) SCA for TCCON and OCO-2. (b) Predicted SCA from multiple-linear regression with mean potential temperature at 700 hPa in March and maximum annual SIF. (c) SCA from TCCON site minus SCA for corresponding OCO-2 grid cell. (d) Predicted SCA minus observed SCA.

We also analyzed correlations between the SCA, equivalent latitude (potential temperature at 700 hPa), and maximum annual SIF for the OCO-2 5 degree by 5 degree northern hemisphere grid and created a multiple linear regression model for SCA. Figure 3.6 shows that there exists a moderately strong correlation between SCA and mean March potential temperature at 700 hPa. With an r-squared value of 0.57 for OCO-2 and 0.59 for TCCON, over half of the variation seen between SCAs calculated over small spatial scales can be explained by mean March potential temperature.

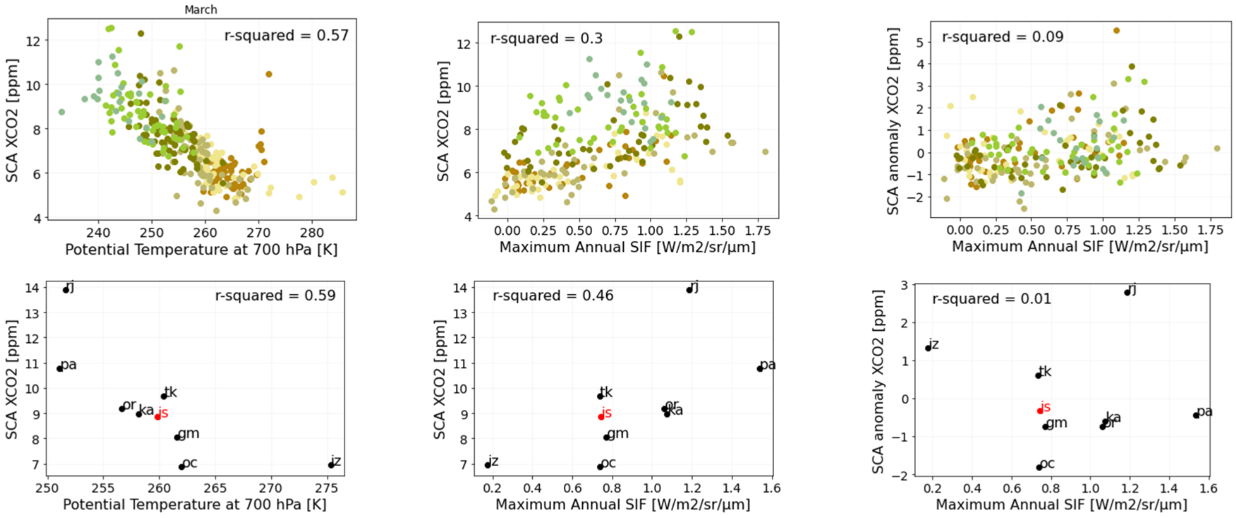


Figure 3.6 SCA from OCO-2 correlated with (a) mean potential temperature at 700 hPa in March and (b) maximum annual SIF, colored by latitude. SCA from TCCON correlated with (d) mean potential temperature at 700 hPa in March and (e) mean diurnal cycle amplitude in December with urban sites colored red. SCA anomaly correlated with maximum annual SIF for (c) OCO-2 and (f) TCCON.

To determine the contribution of ecosystem productivity on SCA at a given location, we analyzed the relationship between SCA and maximum annual SIF calculated at a given location (Fig. 3.6). We see weaker correlations, and the r-squared value of 0.30 for OCO-2 and 0.46 for TCCON indicates that at least 30% of the SCA measured at a location is explained by SIF within the same area. However, the maximum annual SIF does not explain much of the SCA anomaly, calculated as the difference between the observed SCA and the predicted SCA at that equivalent latitude. Completing multiple linear regression with equivalent latitude and maximum annual SIF shows a 10% increase in ability to predict the SCA over using equivalent latitude alone, with an r-squared of 0.62 for both OCO-2 and TCCON.

As a test of the impact of local exchange on the total column averaged SCA in a given location, we analyzed the correlation between SCA and DCA for each TCCON site (Fig. 3.7). We see the largest correlation for December. The r-squared value of 0.35 indicates that over 30%

of the variation in SCA between TCCON sites can be explained by local exchange. DCA is able to explain some of the variation in the SCA anomaly and we see an r-squared of 0.30 when relating SCA anomaly to mean December DCA. There is a substantial 20% improvement in ability to predict SCA using multiple linear regression with equivalent latitude and mean December DCA compared to using equivalent latitude alone.

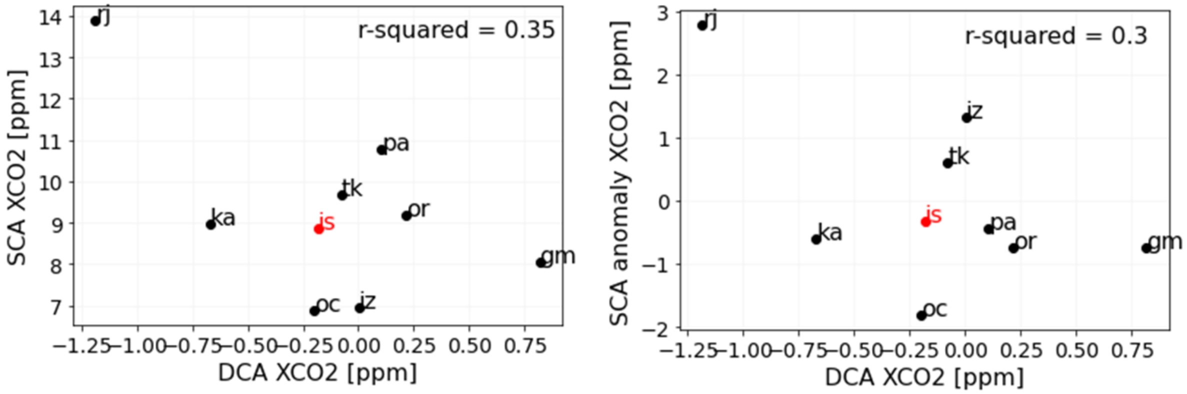


Figure 3.7 (a) SCA from TCCON correlated with mean diurnal cycle amplitude in December with urban sites colored red. (b) SCA anomaly from TCCON correlated with mean diurnal cycle amplitude in December with urban sites colored red.

We also compared how the predicted SCA using the multiple linear regression results at a location compared to the observed SCA for OCO-2 (Fig. 3.5 b,d) and TCCON (Fig. 3.8). For OCO-2, we are only able to compare over land, where SIF occurs. The multiple linear regression model is able to reproduce much of the spatial distribution on land, aside from Equatorial Africa where observed SCA are nearly twice as large as the results using multiple linear regression. In general, MLR both underestimates and overestimates the SCA, but the largest differences are underestimates. In addition to Equatorial Africa, we also see significant underestimates in India and Northeastern Asia.

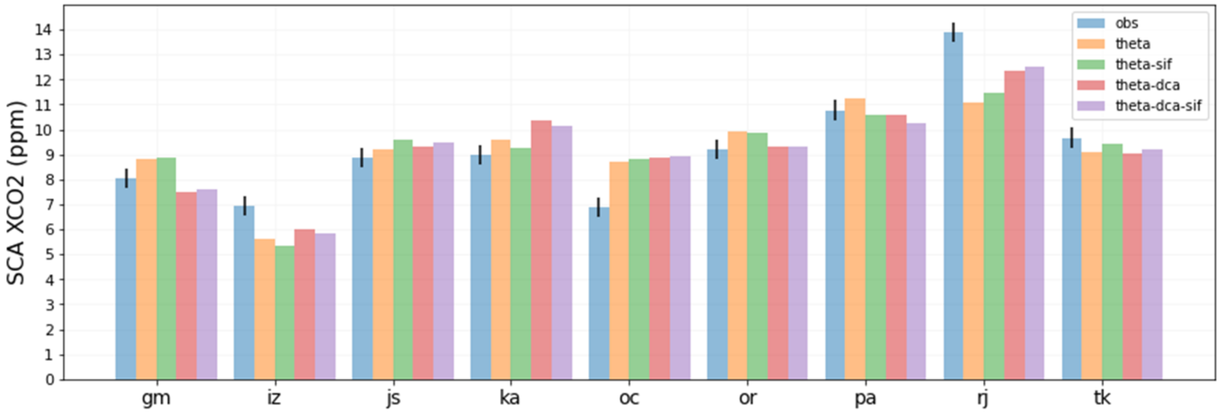


Figure 3.8 Observed and predicted SCA at TCCON sites.

Similarly to OCO-2, for TCCON we see that while the prediction may overestimate or underestimate the SCA, and the largest differences are seen when the predicted SCA is smaller than the observed. However, the SCA is overestimated at the majority of TCCON sites but was overestimated and underestimated over roughly equal areas of the map for OCO-2. This is likely because most of the TCCON sites included in this analysis lies between 30 ° N and 60 ° N, where the OCO-2 SCA is also generally smaller than predicted. The exception in this region is Northeastern Asia, where there is a large underestimate of the SCA relative to observations. This is also true for TCCON, where the Rikubetsu, Japan site also shows a large underestimate of the SCA relative to observations.

There is no significant improvement in the estimate when including equivalent latitude, mean December DCA, and maximum annual SIF over using equivalent latitude with just SIF or just DCA. In general, multiple linear regression with mean December DCA and equivalent latitude is able to change the prediction more relative to single linear regression with equivalent latitude than multiple linear regression with maximum annual SIF and equivalent latitude.

3.4 Discussion

3.4.1 Hemispheric Seasonal Cycle

The seasonal cycle obtained by taking the area-weighted average of the data retrieved from TCCON sites between 25° N and 55° N compares favorably with the seasonal cycle obtained using OCO-2 observations. The shape, timing, and amplitude of the seasonal cycles for TCCON and OCO-2 agree reasonably well with each other and the SCA of 6.9 ± 0.4 ppm for TCCON is just 0.1 ppm higher than the SCA of 6.8 ± 0.1 ppm for OCO-2 (Fig. 3.3). The agreement between the spatially dense but temporally sparse OCO-2 measurements and the spatially sparse but temporally dense TCCON measurements indicates that the network may be able to be representative of the latitude band analyzed. A recent study by Labzovskii et al. (2021) showed the promise of TCCON to capture the global CO_2 growth rate. Labzovskii et al. (2021) found the estimated global Annual Growth Rate of CO_2 using aggregated TCCON measurements was independent from differences in the time series between TCCON stations and the location of the stations used relative to neighboring CO_2 sources. The consistency between the TCCON and OCO-2 northern seasonal cycles, along with the results of Labzovskii et al. (2021), highlights the potential usefulness of TCCON in monitoring and understanding carbon exchange in northern ecosystems.

A note of caution is due here because while TCCON and OCO-2 agree well with each other, neither agree with the cycle found using measurements from the HIPPO and ATom campaigns. After adjusting for stratospheric contribution to the seasonal cycle, the adjusted-TCCON SCA of 8.4 ± 0.3 ppm is still at least 0.7 ppm lower than the HIPPO and ATom SCA. The method of defining the stratosphere has some effect on the adjusted-SCA value inferred from the TCCON retrievals. The boundary of 300 hPa was chosen to be consistent with the methodology applied to the HIPPO and ATom observations. Moving this boundary up to 250 hPa decreases the adjusted-SCA to 8.3 ppm and moving the boundary down to 350 hPa increases the adjusted-SCA to 8.7 ppm. Additionally, the tropopause is defined in the TCCON prior profiles. When using the tropopause altitude to define the cutoff, the adjusted-SCA decreases to 8.3 ppm. All of these methods results in a SCA that falls within the uncertainty bounds and none of these methods brings TCCON into agreement with HIPPO and ATom.

Furthermore, the analysis using CT2019B revealed that when sampled like TCCON, the SCA is 0.7 ppm too low when considering the total column and over 1 ppm too low when considering the tropospheric column. This explains the mismatch between the adjusted-TCCON seasonal cycle and the HIPPO and ATom seasonal cycle and indicates that TCCON is not yet extensive enough to fully capture the average northern hemisphere seasonal cycle.

3.4.2 Site-to-Site Seasonal Cycles

Site-to-site differences in the seasonal cycle of atmospheric carbon can largely be explained by the location of the site. It is known that the SCA of atmospheric CO_2 is positively correlated with latitude in the Northern Hemisphere because of increased seasonality in light availability, temperature, and photosynthesis at higher latitudes (Barnes et al., 2016; Forkel et al., 2016; Jacobs et al., 2021; Randerson et al., 1997), giving latitude some predictive power on SCA. We see a moderately strong correlation between SCA of atmospheric CO_2 and equivalent latitude for the northern TCCON sites evaluated ($r^2 = 0.59$, $p < 0.05$) as well as for OCO-2 ($r^2 = 0.57$, $p < 0.05$), consistent with these prior results (Fig. 3.6). This indicates that roughly half of the variation seen in SCA can be explained by equivalent latitude. However, this relationship is not as strong in the mid-latitudes indicating that there are other sources of variation present. For example, Jacobs et al. (2021) found significant longitudinal gradients in SCA spanning latitudes from 47 °N to 72 °N driven primarily by atmospheric transport. We also see significant differences at small spatial scales. Five sites in Europe (Bremen, Germany; Karlsruhe, Germany; Garmisch, Germany; Orleans, France; and Paris, France) are all contained within a circle of radius 500 km and yet the SCA measured at these sites ranges from 7.9 ppm in Garmisch to 10.2 ppm in Paris.

Local ecosystem productivity will impact the SCA measured at any location and SIF can be used as a proxy for local productivity as it has been shown to correlate with GPP (e.g. Frankenberg et al., 2011; Guanter et al., 2014; Li et al., 2018; Magney et al., 2019). The correlation between the SCA and the maximum annual SIF can be used to explain around 30% of variation in OCO-2 SCA (Fig. 3.6) and 46% of variation in TCCON SCA. Taking correlations with equivalent latitude and SIF, 62% of the variation in SCA can be explained.

The mean diurnal cycle amplitude (DCA) in December of each TCCON site is also correlated with the SCA. With an r-squared value of 0.35 and a p-value less than 0.05, it is revealed that over 30% of the site-to-site variation can be explained by the diurnal cycle (Fig. 3.6). The diurnal cycle occurs on a short timescale and the diurnal flux in a region has been shown to depend on local NEE (Zhang et al., 2015) and seasonal circulation (Metya et al., 2021). The DCA improves the predicted SCA when applying a multiple linear regression compared to using SIF, indicating the importance of NEE and seasonal circulation in determining the SCA at a given location. This is interesting in light of the anomalously large SCA at the Rikubetsu, Japan TCCON site, thought to be largely due to atmospheric transport, being accompanied by a large DCA.

We saw that using multiple linear regression results to predict SCA for OCO-2 and TCCON underestimates the SCA in Equatorial Africa, Northeastern Asia, and India. Palmer et al., (2019) also see an unexpectedly large CO_2 seasonal cycle over Northern Tropical Africa using satellite data. They saw the largest uptake over the Congo basin, as expected, and also saw large emissions during March and April over Western Ethiopia and Western Tropical Africa. Palmer et al., (2019) were able to rule out systematic errors as a reason for the large carbon seasonal cycle in that region. The large SCA in Northeastern Asia was also seen in Asian Boreal Forests by Jacobs et al., (2021). They conclude that the longitudinal gradient of increasing SCA seen as you move eastward across Europe and Asia highlight the importance of global atmospheric transport in determining spatial patterns in XCO_2 seasonal variations. Tiwari et al. (2012) extensively studied variation in atmospheric carbon dioxide over India. They saw that the seasonal cycle is heavily affected by rainfall during the monsoon seasons, showing both a summer and winter minimum that correlates with the summer and winter monsoon. However, it is unclear if this double minimum would lead to a larger SCA over the region.

3.5 Conclusions

This study set out to answer the questions:

1. Can remote sensing observations of XCO_2 provide a hemispheric constraint that is consistent with that from HIPPO and ATom?
2. Do we understand site-level and grid scale-level patterns in seasonal carbon exchange?

In response to the first question, comparison of the seasonal cycle found using TCCON retrievals to those found using OCO-2 retrievals and HIPPO and ATom measurements revealed that while the existing TCCON is consistent with the OCO-2 zonal mean, the TCCON network biased low in comparison to HIPPO and ATom measurements. In response to the second question, linear regression analysis revealed that potential temperature, diurnal cycle, and SIF are all able to explain some of the spatial variation in seasonal land-atmosphere carbon exchange.

3.6 References

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), n/a-n/a. https://doi.org/10.1029/2004GB002439

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., et al. 2001. FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society, 82, 2415-2434. https://doi.org/10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2

Baldocchi, D. (2003). Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology*, *9*(4), 479–492. https://doi.org/10.1046/j.1365-2486.2003.00629.x

Barnes, E. A., Parazoo, N., Orbe, C., & Denning, A. S. (2016). Isentropic transport and the seasonal cycle amplitude of CO $_2$. *Journal of Geophysical Research: Atmospheres*, *121*(13), 8106–8124. https://doi.org/10.1002/2016JD025109

Blumenstock, T., F. Hase, M. Schneider, O.E. García, E. Sepúlveda. 2017. TCCON data from Izana, Tenerife, Spain, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.izana01.R1

Chevallier, F., Deutscher, N. M., Conway, T. J., Ciais, P., Ciattaglia, L., Dohe, S., et al. (2011). Global CO $_2$ fluxes inferred from surface air-sample measurements and from TCCON retrievals of the CO $_2$ total column. *Geophysical Research Letters*, *38*(24), n/a-n/a. https://doi.org/10.1029/2011GL049899

Ciais, P., Tan, J., Wang, X., Roedenbeck, C., Chevallier, F., Piao, S. L., Moriarty, R., et al. (2019). Five decades of northern land carbon uptake revealed by the interhemispheric CO2 gradient. *Nature*, *568*(7751), 221–225. https://doi.org/10.1038/s41586-019-1078-6

Cleveland, R. B., Cleveland, W. S., McRae, J. E., & Terpenning, I. (1990). STL: A seasonal-trend decomposition. *J. Off. Stat*, *6*(1), 3-73.

Eldering, A., O'Dell, C. W., Wennberg, P. O., Crisp, D., Gunson, M. R., Viatte, C., et al. The Orbiting Carbon Observatory-2: first 18 months of science data products, Atmos. Meas. Tech., 10, (2017) 549–563, https://doi.org/10.5194/amt-10-549-2017.

Erickson, D.J., III, P.J. Rasch, P.P. Tans, P. Friedlingstein, P. Ciais, E. Maier-Reimer, et al. The seasonal cycle of atmospheric CO2: A study based on the NCAR Community Climate Model (CCM2), J. Geophys. Res., 101(DlO), 15,079-15,097, 1996.

Frankenberg, C., Fisher, J. B., Worden, J., Badgley, G., Saatchi, S. S., Lee, J.-E., et al. (2011). New global observations of the terrestrial carbon cycle from GOSAT: Patterns of plant fluorescence with gross primary productivity. *Geophysical Research Letters*, *38*(17), n/a-n/a. https://doi.org/10.1029/2011GL048738

Forkel, M., Carvalhais, N., Rödenbeck, C., Keeling, R., Heimann, M., Thonicke, K., et al. (2016). Enhanced seasonal CO $_2$ exchange caused by amplified plant productivity in northern ecosystems. *Science*, *351*(6274), 696–699. https://doi.org/10.1126/science.aac4971

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., et al. (2013). Enhanced Seasonal Exchange of CO2 by Northern Ecosystems Since 1960. *Science*, *341*(6150). https://doi.org/10.1126/science.1239207

Guanter, L., Zhang, Y., Jung, M., Joiner, J., Voigt, M., Berry, J. A., ... & Griffis, T. J. (2014). Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. *Proceedings of the National Academy of Sciences*, *111*(14), E1327-E1333.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., et al. (2003). TransCom 3 CO $_2$ inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information. *Tellus B: Chemical and Physical Meteorology*, *55*(2), 555–579. https://doi.org/10.3402/tellusb.v55i2.16728

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., et al. (2004). Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochemical Cycles*, *18*(1), n/a-n/a. https://doi.org/10.1029/2003GB002111

Hase, F., T. Blumenstock, S. Dohe, J. Groß, M. Kiel. 2017. TCCON data from Karlsruhe, Germany, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.karlsruhe01.R0

Jacobs, N., Simpson, W. R., Graham, K. A., Holmes, C., Hase, F., Blumenstock, T., et al. (n.d.). *Spatial distributions of X CO 2 seasonal cycle amplitude and phase over northern high latitude regions*. https://doi.org/10.5194/acp-2021-185

Jung, M., Koirala, S., Weber, U. *et al.* The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Sci Data* **6**, 74 (2019). https://doi.org/10.1038/s41597-019-0076-8

Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, *17*(5), 1343–1365. https://doi.org/10.5194/bg-17-1343-2020

Keppel-Aleks, G., Wennberg, P. O., Washenfelder, R. A., Wunch, D., Schneider, T., et al. (2012). The imprint of surface fluxes and transport on variations in total column carbon dioxide. *Biogeosciences*, *9*(3), 875–891. https://doi.org/10.5194/bg-9-875-2012

Kivi, R., P. Heikkinen, E. Kyro. 2017. TCCON data from Sodankyla, Finland, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.sodankyla01.R0

Knorr, W., and M. Heimann, Impact of drought stress and other factors on seasonal land biosphere CO 2 exchange studied through an atmospheric tracer transport model, Tellus, Ser. B, 47, 471-489, 1995.

Labzovskii, Lev D., Samuel Takele Kenea, Hannakaisa Lindqvist, Jinwon Kim, et al. 2021. "Towards Robust Calculation of Interannual CO_2 Growth Signal from TCCON (Total

Carbon Column Observing Network)" *Remote Sensing* 13, no. 19: 3868. https://doi.org/10.3390/rs13193868

Laughner, J. L.**,** G. C. Toon, D. Wunch, S. Roche, J. Mendonca, M. Kiel, C. M. Roehl, et al. "The Total Carbon Column Observing Network's GGG2020 Data Version." *Earth Sys. Sci. Data*, (in prep)

Magney, Troy S., David R. Bowling, Barry A. Logan, Katja Grossmann, Jochen Stutz, et al.. "Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence." *Proceedings of the National Academy of Sciences* 116, no. 24 (2019): 11640-11645.

Metya, A., Datye, A., Chakraborty, S. *et al.* Diurnal and seasonal variability of CO_2 and CH_4 concentration in a semi-urban environment of western India. *Sci Rep* **11**, 2931 (2021). https://doi.org/10.1038/s41598-021-82321-1

Morino, I., V. A. Velazco, A. Hori, O. Uchino, D. W. T. Griffith. 2017a. TCCON data from Burgos, Philippines, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.burgos01.R0

Morino, I., H. Ohyama, A. Hori, H. Ikegami. 2017b. TCCON data from Rikubetsu, Hokkaido, Japan, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.rikubetsu01.R0

Morino, I., H. Ohyama, A. Hori, H. Ikegami. 2017c. TCCON data from Tsukuba, Ibaraki, Japan, 125HR, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. http://doi.org/10.14291/tccon.ggg2020.tsukuba02.R0

Notholt, J., C. Petri, T. Warneke, N. Deutscher, M. Buschmann, C. Weinzierl, R. Macatangay, P. Grupe. 2019. TCCON data from Bremen, Germany, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.bremen01.R0

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., et al. Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018

Palmer, P.I., Feng, L., Baker, D. *et al.* Net carbon emissions from African biosphere dominate pan-tropical atmospheric CO_2 signal. *Nat Commun* **10**, 3344 (2019). https://doi.org/10.1038/s41467-019-11097-w

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

Randerson, J. T., Thompson, M. v., Conway, T. J., Fung, I. Y., & Field, C. B. (1997). The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochemical Cycles*, *11*(4), 535–560. https://doi.org/10.1029/97GB02268

Shiomi, K., S. Kawakami, H. Ohyama, K. Arai, H. Okumura, H. Ikegami, M. Usami. 2022. TCCON data from Saga, Japan, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.saga01.R0

Sun, Y., Frankenberg, C., Jung, M., Joiner, J., Guanter, L., Köhler, P., & Magney, T. (2018). Overview of Solar-Induced chlorophyll Fluorescence (SIF) from the Orbiting Carbon Observatory-2: Retrieval, cross-mission comparison, and global monitoring for GPP. *Remote Sensing of Environment*, *209*, 808-823. https://doi.org/10.1016/j.rse.2018.02.016

Sussmann, R., M. Rettinger. 2017. TCCON data from Garmisch, Germany, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.garmisch01.R0

Total Carbon Column Observing Network (TCCON) Team. (2022). 2020 TCCON Data Release (Version GGG2020) [Data set]. CaltechDATA. https://doi.org/10.14291/TCCON.GGG2020

Te, Y., P. Jeseck, C. Janssen. 2017. TCCON data from Paris, France, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.paris01.R0

Tucker, C.J., I.Y. Fung, C.D. Keeling, and R.H. Gammon, 1986: Relationship between atmospheric CO_2 variations and a satellite-derived vegetation index. *Nature*, **319**, 195-199, doi:10.1038/319195a0.

Warneke, T., J. Messerschmidt, J. Notholt, C. Weinzierl, N. Deutscher, C. Petri, et al. 2017. TCCON data from Orleans, France, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.orleans01.R0

Washenfelder, R. A., Toon, G. C., Blavier, J.-F., Yang, Z., Allen, N. T., Wennberg, P. O., et al. (2006), Carbon dioxide column abundances at the Wisconsin Tall Tower site, *J. Geophys. Res.*, 111, D22305, doi:10.1029/2006JD007154.

Wennberg, P. O., D. Wunch, C. Roehl , J.-F. Blavier, G. C. Toon, N. Allen. 2017a. TCCON data from California Institute of Technology, Pasadena, California, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.pasadena01.R0

Wennberg, P. O., D. Wunch, C. Roehl, J.-F. Blavier, G. C. Toon, N. Allen, P. Dowell, et al. 2017b. TCCON data from Lamont, Oklahoma, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.lamont01.R0

Wennberg, P. O., C. Roehl, D. Wunch, G. C. Toon, J.-F. Blavier, R. Washenfelder, et al. 2017c. TCCON data from Park Falls, Wisconsin, USA, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. http://doi.org/10.14291/tccon.ggg2020.parkfalls01.R0

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., & Yavin, Y. G. (2009). Emissions of greenhouse gases from a North American megacity. *Geophysical Research Letters*, *36*(15), n/a-n/a. https://doi.org/10.1029/2009GL039825

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., et al. Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, (2010) 1351-1362; https://doi:10.5194/amt-3-1351-2010

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., et al. (2011). The Total Carbon Column Observing Network. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1943), 2087–2112. https://doi.org/10.1098/rsta.2010.0240

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., et al.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) X_{CO2} measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017

Wunch, D., J. Mendonca, O. Colebatch, N. Allen, J.-F. L. Blavier, K. Kunz, et al. 2020. TCCON data from East Trout Lake, Canada, Release GGG2020R0. TCCON data archive, hosted by CaltechDATA, California Institute of Technology, Pasadena, CA, U.S.A. https://doi.org/10.14291/tccon.ggg2020.easttroutlake01.R0

3.7 Supplement

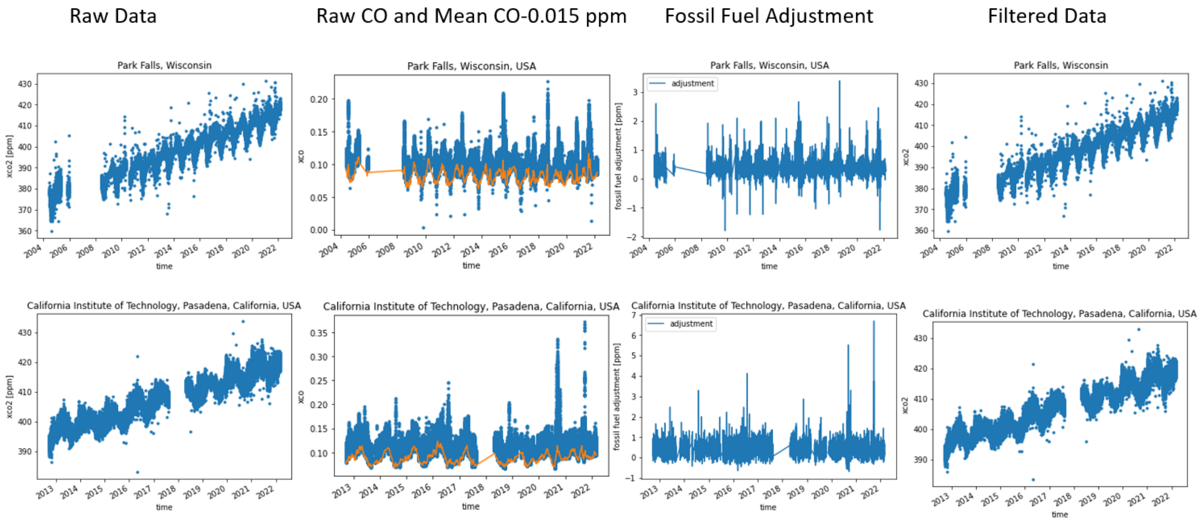


Figure 3.9 (S1) Fossil fuel correction of TCCON data at Park Falls, Wisconsin (forested site) and Pasadena, California (urban site).

Chapter 4 **Conclusion**

The goal of the work described here was to address large uncertainties in model simulation and land flux estimates by leveraging measurements of atmospheric CO_2 to evaluate model performance and increase understanding of the spatial distribution of land carbon fluxes. The first aim was to estimate the Northern extratropical GSNF using aircraft profiles of CO_2 measured over the remote oceans and use the GSNF as a model benchmark. The results indicate that the GSNF for the northern extratropics is 5.7 ± 0.3 Pg C and that CMIP5 and CMIP6 models tend to underestimate the GSNF and overestimate the growing season length. This result provides a new and large-scale observational target for prognostic model evaluation. The second aim was to estimate northern extratropical annual fluxes for GPP, RH, and NPP using an emergent constraint approach applied to prognostic models. Our GSNF-constrained value for GPP is 56 ± 15 Pg C, which is 8 Pg C larger than a commonly used estimate derived from flux towers. Given the discrepancy between the hemispheric GSNF value and the upscaled value from flux towers – which provide very local information – we sought to determine constraints on growing season fluxes at intermediate spatial scales.

To do this, we defined our third aim, to upscale the TCCON network to estimate the northern hemisphere seasonal cycle. The results indicate that the TCCON network may be representative of certain latitude bands, but is not representative of the northern hemisphere as a whole. Increasing the coverage of the network, especially at high latitudes, may improve the representativeness of the network. We also delved into diurnal cycle information from TCCON sites and spatially resolved variations in productivity from space-based data to understand spatial heterogeneity in seasonal amplitudes. We saw a moderate correlation between the seasonal cycle amplitude and equivalent latitude indicating that approximately 60% of the seasonal cycle at a given site can be explained by the equivalent latitude of that site, meaning it reflects large-scale fluxes. The correlation between seasonal cycle amplitude anomaly and diurnal cycle amplitude indicates that the diurnal cycle can explain about 30% of the remaining difference between sites, meaning that the imprint of local fluxes can also be derived from observations atmospheric CO2, which had previously been shown to contain mostly large-scale information. The results of multiple linear regression indicated that potential temperature and SIF can explain approximately 60% of the variation and potential temperature and DCA can explain approximately 70% of the variation. In addition to providing insights about the large-scale versus local information content of the observations, our approach enables us to identify locations that our simple model does not fit well. The multiple linear regression fails to accurately predict the SCA in Equatorial Africa, Northeastern Asia, and India. These area of ddisagreement may provide a roadmap for areas where future carbon cycle studies should focus.

All of the findings discussed here were subject to some limitations. The GSNF inferred using HIPPO and ATom data relies upon inverse model output, which is a large source of uncertainty in the inferred flux. The constrained values for GPP, RH, and NPP have large uncertainties due to the difficulty of directly measuring these fluxes at large spatial scales. Retrievals of column-averaged CO_2 are limited to summer months in the high latitudes when sunlight is available, making it difficult to fully constrain the seasonal cycle at high latitudes. This makes it difficult to gain hemispheric level information from these types of observations. Additionally, correlation analysis can only show the effect of the variables being explored and there may be additional explanatory variables that were not included in the analysis, such as heterotrophic respiration, which is not currently observed via remote sensing. Despite these limitations and others not explicitly stated here, this work shows the power of high-quality atmospheric CO_2 observations to shed light on the magnitude and spatial distribution of northern land carbon fluxes.

Appendix

Appendix A - Acronyms and Abbreviations

ATom – Atmospheric Tomography Mission

CMIP – Coupled Model Intercomparison Project

CO$_2$ – carbon dioxide

GHG – Greenhouse Gas

GPP – gross primary productivity

GSNF – growing season net flux

HIAPER – High-performance Instrumented Airborne Platform for Environmental Research

HIPPO – HIAPER Pole-to-Pole Observations

ILAMB – International Land Model Benchmarking

NBP – net biosphere productivity

NPP – net primary productivity

OCO-2 – Orbiting Carbon Observatory 2

RH – heterotrophic respiration

SIF – solar induced chlorophyll fluorescence

TCCON – Total Carbon Column Observing Network

XCO$_2$ – column-average dry air mole-fraction of carbon dioxide