CarbonTracker Documentation
CT2026 release

September 18, 2026

Authors: CarbonTracker CT2026 Team
Full list of authors and their affiliations in Appendix A

Contents

1 Introduction **3**
1.1 A tool for science, and for policy . . . 3
1.2 The role of other atmospheric species in constraining the atmospheric carbon budget . . . 4
1.3 Updates . . . 4
1.4 Citation and usage policy . . . 5
1.4.1 Usage Policy . . . 5
1.4.2 Citing our results . . . 5

2 Model description **7**

3 Prior models **9**
3.1 Fossil fuel module . . . 9
3.1.1 The "Miller" emissions dataset . . . 10
3.1.2 The "OCO-2" emissions dataset . . . 12
3.1.3 Fossil fuel emissions datasets comparisons . . . 13
3.1.4 Uncertainties . . . 13
3.2 Terrestrial biosphere module . . . 14
3.2.1 MiCASA model . . . 14
3.2.2 Temporal downscaling . . . 16
3.2.3 SiB4 . . . 17
3.2.4 Comparison of terrestrial biosphere priors . . . 18
3.3 Fire module . . . 18
3.3.1 GFED fire emissions model . . . 20
3.3.2 The MiCASA GFED v3 variant . . . 20
3.3.3 GFED4.1s fire emissions . . . 20
3.3.4 Fire emissions comparison . . . 20
3.4 Oceans module . . . 22
3.4.1 Air-sea CO_2 exchange . . . 23
3.4.2 SOM-FFN model . . . 23
3.4.3 AOML-Extra Trees model . . . 23
3.4.4 Gas-transfer velocity and ocean surface properties . . . 24

3.4.5 Specifics of the inversion methodology related to air-sea CO_2 fluxes . . . 25
3.4.6 Prior ocean emissions comparison . . . 25

4 Atmospheric transport **26**
4.1 TM5 offline tracer transport model . . . 26
4.2 Improvement in convective transport . . . 27

5 Observations **30**
5.1 The CarbonTracker observational network . . . 30
5.2 Adaptive model-data mismatch . . . 34

6 Ensemble data assimilation **36**
6.1 Parameterization of unknowns . . . 36
6.1.1 Optimization regions . . . 37
6.1.2 Assimilation window . . . 37
6.2 Physical parallelization . . . 39
6.3 Dynamical model . . . 40
6.3.1 Structure of master prior covariance . . . 41
6.3.2 Process Noise . . . 41
6.3.3 Posterior uncertainties in CarbonTracker . . . 43

7 Statistical performance of CT2026 **44**
7.1 Measurement data . . . 44

8 Resources and References **46**

A Full author list **55**

B Performance by dataset **60**

C Ecoregions in CarbonTracker **102**
C.1 What are ecoregions? . . . 102
C.2 Why use ecoregions? . . . 103
C.3 Ecosystems within Transcom regions . . . 103

Chapter 1

Introduction

The goal of the CarbonTracker program is to produce quantitative estimates of atmospheric carbon uptake and release at the Earth's surface that are consistent with observed patterns of CO_2 in the atmosphere. CarbonTracker is an *inverse* model of atmospheric CO_2, which means that it attempts to match atmospheric CO_2 measurements by adjusting inputs and removals of carbon dioxide at the Earth's surface until they best agree with CO_2 measurements and atmospheric transport.

CarbonTracker is updated on an approximately-annual basis. The current release, CT2026, provides results from 2000 through the end of 2025. A "near-real" time model product, CT-NRT, extends these results after the most recent CarbonTracker release. Available model versions are listed at https://gml.noaa.gov/ccgg/carbontracker/version.php

1.1 A tool for science, and for policy

CarbonTracker is made possible by the long-term monitoring of atmospheric CO_2 conducted by many academic and governmental programs around the world (see Section 5). These data help improve our understanding of how the land and ocean are responding to Earth's changing climate. The uptake and release of CO_2 by these ecosystems is changing due to chemical and physical responses to increased atmospheric CO_2 concentrations, to human management of lands and oceans, and to changes in temperature, precipitation, and winds.

CarbonTracker is a completely open product. All results, including graphics and tabular data, may be freely used without restriction, although we do request the favor of appropriate acknowledgment (see Section 1.4 and https://gml.noaa.gov/ccgg/carbontracker/citation.php).

The unrestricted access to all CarbonTracker results means that anyone can scrutinize our work, suggest improvements, and profit from our efforts. We hope this scrutiny will help guide further development of our methods, and improve our ability to monitor, diagnose, and possibly predict the behavior of the global carbon cycle. We encourage collaborations focused on use of the CarbonTracker model as a tool for scientific analysis. Please contact us if you would like to get involved and collaborate with us.

CarbonTracker also can be relevant for helping to inform carbon policy. Its ability to accurately quantify natural and anthropogenic emissions and uptake at regional scales is currently limited by a sparse observational network. With enough observations however, CarbonTracker and systems like it will be able to monitor regional emissions, including those from fossil fuel use. This will provide an independent check on emissions accounting, including estimates of fossil fuel use based on economic inventories. It can thus

provide feedback to policies aimed at limiting greenhouse gas emissions. This independent evaluation of the effectiveness of carbon policy is the bottom line in any mitigation strategy. It has the added advantage of being a constraint provided by the atmosphere itself, where CO_2 levels matter most.

1.2 The role of other atmospheric species in constraining the atmospheric carbon budget

Many laboratories making high accuracy CO_2 observations also make many other measurements of the same air, typically other greenhouse gases such as methane (CH_4), nitrous oxide (N_2O), sulfur hexafluoride (SF_6), as well as carbon monoxide (CO) and isotopic ratios of CO_2 and CH_4. These measurements are usually reported as mole fractions, for reasons explained here.

These trace gases are relevant for the study of climate change and interesting in their own right, but the additional measurements can also help in identifying sources and sinks of carbon or in understanding carbon cycle processes. For this reason, many air samples are now analyzed for a suite of halocompounds and hydrocarbons. Several of these species can be useful for monitoring air quality, but they can also help with better source apportionment of the greenhouse gases. In addition, the estimation of the source strengths of a number of pollutants could be greatly improved if we were able to quantify fossil fuel CO_2 emissions from air measurements for specified regions.

The best tracer for quantifying the component of atmospheric CO_2 that has been recently added to an air mass through the burning of fossil fuels is the carbon-14 (^{14}C) content of CO_2. Cosmic rays produce ^{14}C, a radioactive form of carbon, in the higher regions of the atmosphere. It is present in the atmosphere and oceans and in all living organisms and their remains, but coal, oil, and natural gas contain no ^{14}C because it has long decayed away. Currently, $^{14}CO_2$ measurements are made on only a small subset of the air samples because of higher analysis costs. None of these other data and their relationships have been used directly in this release of CarbonTracker. We expect them to be incorporated incrementally at later stages.

CarbonTracker is a NOAA contribution to the North American Carbon Program.

1.3 Updates

CarbonTracker is updated about once per year to include new data and model improvements. CT2026 provides results from 2000 through 2025. Previous versions of CarbonTracker and our CT-NRT (CarbonTracker Near-Real Time) releases are available at the CarbonTracker website.

Important revisions of our methods for CT2026 include the following:

- Extension through the end of 2025,
- Revision of fossil fuel emissions,
- New land and wildfire priors,
- Revised air-sea gas exchange including two new pCO_2 prior models,
- A new Kalman filter time-propagation model, and
- Physical parallelization of the analysis period.

1.4 Citation and usage policy

1.4.1 Usage Policy

CarbonTracker is an open product of NOAA's Global Monitoring Laboratory (GML) using data from the international greenhouse gas observational network. Results, including figures and tabular material found on the CarbonTracker website may be used for non-commercial purposes without restriction. We kindly ask you to acknowledge, cite, and/or reference CarbonTracker as described below.

1.4.2 Citing our results

- We ask that scientific work that relies heavily on CarbonTracker products is discussed with us before publication, to ensure proper representation of our work and co-authorship if appropriate.
- Please cite as Jacobson et al. (2026).
- The DOI for CT2026 and all its associated products and results is http://dx.doi.org/10.25925/hqp0-rk68.
- Please use "CT2026" as the shorthand to refer to our product, not "CT". This identifies both the product and the release version. It is vital to identify the version of the product you are using.
- Note that the product is called "CarbonTracker" without a space character, not "Carbon Tracker".
- Please include our suggested acknowledgment text in your acknowledgments section.
- Boilerplate model description text provided upon request.

Example . . . we compare our results to NOAA's CarbonTracker, version CT2026 (Jacobson et al., 2026). In this work, CT2026 is . . .

Acknowledgments CarbonTracker CT2026 results provided by NOAA GML, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov.

Reference Andrew R. Jacobson, Kenneth N. Schuldt, John B. Miller, Ashley Pera, Aleya Kaushik, Arlyn Andrews, Sourish Basu, Joaquin Triñanes, Peter Landschützer, John Mund, Brad Weir, Lesley Ott, Tuula Aalto, Hermanni Aaltonen, James Brice Abshire, Ken Aikin, Grant Allen, Marcos Andrade, Francesco Apadula, Sabrina Arnold, Bianca Baier, Peter Bakwin, Jakub Bartyzel, Gilles Bentz, Peter Bergamaschi, Andreas Beyersdorf, Tobias Biermann, Sebastien C. Biraud, Pierre-Eric Blanc, Harald Boenisch, David Bowling, Gordon Brailsford, Willi A. Brand, Dominik Brunner, Thao Paul V. Bui, Benoit Burban, Lukas Bäni, Francescopiero Calzolari, Cecilia S. Chang, Gao Chen, Huilin Chen, Lukasz Chmura, Jason M. St. Clair, Shane Clark, Sites Climadat, Julian Della Coletta, Aurelie Colomb, Roisin Commane, Lino Condori, Franz Conen, Sébastien Conil, Cédric Couret, Paolo Cristofanelli, Emilio Cuevas, Roger Curcoll, Bruce Daube, Kenneth J. Davis, Martine De Mazière, Rodrigo A. F. de Souza, Stephan De Wekker, Jonathan M. Dean-Day, Marc Delmotte, Tatiana Di Iorio, Alcide Giorgio di Sarra, Russell Dickerson, Elizabeth DiGangi, Joshua P. DiGangi, Michael Elsasser, Lukas Emmenegger, Shuangxi Fang, Marc L. Fischer, Grant Forster, James France, Arnoud Frumau, Marta Fuente-Lastra, Michal Galkowski, Luciana V. Gatti, Torsten

Gehrlein, Christoph Gerbig, Francois Gheusi, Emanuel Gloor, Daisuke Goto, Tim Griffis, Samuel Hammer, Thomas F. Hanisco, Chad Hanson, Shigeru Hashimoto, László Haszpra, Juha Hatakka, Martin Heimann, Michal Heliasz, Daniela Heltai, Stephan Henne, Arjan Hensen, Christian Hermans, Ove Hermansen, Jack Higgs, Eric Hintsa, Antje Hoheisel, Jutta Holst, Laura T. Iraci, Viktor Ivakhov, Daniel A. Jaffe, Lilian Joly, Armin Jordan, Warren Joubert, Hui-Yun Kang, Anna Karion, Stephan Randolph Kawa, Victor Kazan, Ralph F. Keeling, Ishijima Kentaro, Petri Keronen, Jooil Kim, Jörg Klausen, Tobias Kneuer, Mi-Young Ko, Pasi Kolari, Kateřina Komínková, Eric Kort, Elena Kozlova, Paul Krummel, Dagmar Kubistin, Susan S. Kulawik, Nicolas Kumps, Casper Labuschagne, David H.Y. Lam, Xin Lan, Ray L. Langenfelds, Andrea Lanza, Eric Larmanou, Olivier Laurent, Thomas Lauvaux, Jost Lavric, Beverly E. Law, Choong-Hoon Lee, John Lee, Irene Lehner, Kari Lehtinen, Reimo Leppert, Ari Leskinen, Markus Leuenberger, W.H. Leung, Ingeborg Levin, Janne Levula, John Lin, Matthias Lindauer, Anders Lindroth, Zoe Loh, Morgan Lopez, Timothy J. Lueker, Ingrid T. Luijkx, Chris René Lunder, Toshinobu Machida, Ivan Mammarella, Giovanni Manca, Alistair Manning, Andrew Manning, Michal V. Marek, Per Marklund, Josette E. Marrero, Damien Martin, Melissa Yang Martin, Giordane A. Martins, Hidekazu Matsueda, Anna McAuliffe, Kathryn McKain, Harro Meijer, Frank Meinhardt, Lynne Merchant, Jean-Marc Metzger, N. Mihalopoulos, Natasha L. Miles, Charles E. Miller, Logan Mitchell, Meelis Mölder, Jennifer Müller-Williams, Vanessa Monteiro, Stephen Montzka, Heiko Moossen, Caisa Moreno, Eric Morgan, Josep-Anton Morgui, Shinji Morimoto, Hitoshi Mukai, J. William Munger, David Munro, Mathew Mutuku, Cathrine Lund Myhre, Shin-Ichiro Nakaoka, Jaroslaw Necki, Tim Newberger, Sally Newman, Sylvia Nichol, Euan Nisbet, Yosuke Niwa, David Murithi Njiru, Steffen Manfred Noe, Yukihiro Nojiri, Florian Obersteiner, Simon O'Doherty, Bill Paplawsky, Caroline L. Parworth, Jeff Peischl, Olli Peltola, Wouter Peters, Carole Philippon, Salvatore Piacentino, Jean-Marc Pichon, Penelope Pickers, Steve Piper, Joseph Pitt, Christian Plass-Dülmer, Stephen Matthew Platt, Steve Prinzivalli, Michel Ramonet, Ramon Ramos, Xinrong Ren, Enrique Reyes-Sanchez, Scott J. Richardson, Louis-Jeremy Rigouleau, Haris Riris, Pedro P. Rivas, Michael Rothe, Yves-Alain Roulet, Thomas Ryerson, Ju-Mee Ryoo, Maryann Sargent, Motoki Sasakawa, Bert Scheeren, Martina Schmidt, Tanja Schuck, Marcus Schumacher, Jennifer Seibel, Thomas Seifert, Mahesh Kumar Sha, Paul Shepson, Daegeun Shin, Michael Shook, Christopher D. Sloop, Dan Smale, Gerard Spain, Ann Stavert, David Steger, Martin Steinbacher, Britton Stephens, Colm Sweeney, Lise Lotte Sørensen, Risto Taipale, Shinya Takatsuji, Pieter Tans, Yukio Terao, Kirk Thoning, Helder Timas, Margaret Torn, Pamela Trisolino, Kjetil Tørseth, Jocelyn Turnbull, Pim van den Bulk, Alex Vermeulen, Brian Viner, Gabriela Vitkova, Stephen Walker, Andrew Watson, Ray Weiss, Dietmar Weyrauch, Steven C. Wofsy, Sonja Wolter, Justin Worsey, Doug Worthy, Irène Xueref-Remy, Emma L. Yates, Dickon Young, Camille Yver-Kwok, Sönke Zaehle, Andreas Zahn, Christoph Zellweger, Miroslaw Zimnoch. CarbonTracker CT2026, 2026.DOI: 10.25925/hqp0-rk68

Acknowledgment text "CarbonTracker CT2026 results provided by NOAA GML, Boulder, Colorado, USA from the website at http://carbontracker.noaa.gov."

Suggested Website Citation "CarbonTracker CT2026 http://carbontracker.noaa.gov"

Chapter 2

Model description

CarbonTracker is designed to estimate surface sources and sinks of CO_2 that are consistent with observations of atmospheric CO_2 mole fractions. There are four major components to this effort:

1. Observations are the measurement constraint that we attempt to meet. We assimilate about 3.5 million CO_2 mole fraction measurements collected by international partners between 2000 and 2025, from surface sites, towers, aircraft, and ships. These are discussed in Chapter 5.

2. Atmospheric transport provides the connection between surface fluxes and the measured mole fractions of CO_2. We use the TM5 chemical transport model to take surface fluxes and propagate their signals to measurement sites using reanalysis winds. This model is described in Chapter 4.

3. Prior fluxes are provided by biogeochemical models of surface CO_2 exchange with the atmosphere. If these models and the atmospheric transport we use were perfect, the resulting simulated observations would be in agreement with measured values. However, the match to available observations is not good, suggesting that real fluxes are different from these prior fluxes. CarbonTracker seeks to find the modifications to prior fluxes to yield better agreement with measurements. The collection of prior flux models is discussed in Chapter 3.

   We perform two inversions with independent sets of priors, and the final CT2026 fluxes are the mean of those two inversions. The two inversions are coded as "p7" and "p8" and the priors used in them are listed in Table 2.1.

4. Assimilation is the statistical machinery of finding the optimal fluxes to agree with observations. The unknowns we seek are scaling factors multiplying the prior model CO_2 exchange with the atmosphere for each week over discrete regions. The process of finding the best scaling factors is known as "inverse" modeling. The estimation technique we employ is an ensemble Kalman smoother, as described in Chapter 6.

   The inversion process takes prior fluxes which are not consistent with changes in atmospheric CO_2 mole fraction measurements, and scales them regionally and temporally to meet the observational constraint.

The CT2026 inversion process finds a deeper land sink than that estimated by either of the two prior land models (Table 2.2). There is a large disparity in imposed fire fluxes, a difference of 1.2 Pg C yr^{-1} in the 2001-2025 mean, with the p8 (MiCASA) variant simulating average annual fire emissions of 3.3 Pg C yr^{-1},

Prior set code	Fossil fuels	NEE	Wildfire	Ocean
p7	OCO-2 (Section 3.1.2)	SiB4 (Section 3.2.3)	SiB4/GFED4.1s (Section 3.3.3)	SOM-FFN (Section 3.4.2)
p8	Miller (Section 3.1.1)	MiCASA (Section 3.2.1)	MiCASA (Section 3.3.2)	AOML-ET (Section 3.4.3)

Table 2.1: Prior models used in each CT2026 inversion. "NEE" stands for net ecosystem exchange from the land biosphere, and represents the sum of photosynthetic uptake of CO_2 and release by respiration. Notably, it does not include emissions of CO_2 due to fires.

compared to the smaller p7 (SiB4/GFED4.1s) emissions of 2.1 Pg C yr^{-1}. Since the observational constraint imposed by the inversion process requires that the sum of global fluxes meets the measured growth in atmospheric CO_2, the p8 inversion must find a correspondingly deeper sink in its optimized NEE and ocean fluxes compared to p7. Indeed, we find that in the long-term mean, the sum of NEE plus ocean optimized fluxes for the p8 inversion is -7.7 Pg C yr^{-1}, and that for p7 is just -6.6 Pg C yr^{-1}.

CarbonTracker results can be explored at the CT2026 website, where there are analysis pages on fluxes and CO_2 measurements.

	Prior		Imposed	Posterior			
Model	NEE	Ocean	Fire	NEE	Ocean	NBE	Natural
p7	-4.1	-2.0	2.1	-5.0	-1.6	-2.9	-4.5
p8	-2.4	-1.8	3.3	-5.7	-2.0	-2.4	-4.4
CT2026	-3.2	-1.9	2.7	-5.3	-1.8	-2.6	-4.3

Table 2.2: Long-term mean (2001-2025) prior and posterior global fluxes in the p7 and p8 inversions, and the final CT2026 results. All values are in Pg C yr^{-1}. "NEE" is the non-fire net ecosystem exchange from the land biosphere. Net biosphere exchange ("NBE") is the sum of fire emissions and NEE. "Natural" represents the sum of NBE and ocean fluxes.

Chapter 3

Prior models

CarbonTracker CT2026 is a full reanalysis of the 2000-2025 period using new fossil fuel emissions, new air-sea CO_2 exchange, and new terrestrial biosphere priors. In an attempt to quantify the impact of prior model choice on our final results, we conducted two independent inversions with different sets of prior models. These two inversions were coded as "p7" and "p8". The prior models used for each of these inversions, and the sections of this documents where each is described, are listed in Table 2.1.

3.1 Fossil fuel module

Human beings first influenced the carbon cycle through land-use change. Early humans used fire to control animals and later cleared forests for agriculture. Over the last two centuries, following the industrial and technical revolutions and continuing global population increase, fossil fuel combustion has become the largest anthropogenic source of CO_2. Coal, oil and natural gas combustion are the most common energy sources in both developed and developing countries. Global cement production is also significant, contributing about 5% of total fossil CO_2 emissions. Important sectors of the economy—power generation, transportation, residential & commercial building heating, and industrial processes—rely on fossil fuels. The continued growth of fossil fuel combustion has led to a steady increase of global CO_2 emissions to the atmosphere (Figure 3.1). According to Boden et al. (2017), global emissions of CO_2 from fossil fuel burning, cement manufacturing, and flaring reached 5 billion metric tons of carbon per year (Pg C yr^{-1}) in the decade of the 1970s. Updated emissions products indicate that global total emissions exceeded 10 Pg C yr^{-1} for the first time in 2018. One petagram of carbon, Pg C, is equal to 10^{15} grams of carbon, or one billion metric tons of carbon. To convert to mass of CO_2 emitted, one would multiply by the factor $\frac{44}{12}$, representing the molecular weight of CO_2 compared to the atomic weight of carbon.

U.S. input of CO_2 to the atmosphere from fossil fuel burning in 2025 was 1.4 Pg C, representing 13% of the global total. North American emissions remained nearly constant from 2000-2018, and decreased slightly during the 2020 COVID-19 pandemic year. On the other hand, emissions from developing economies such as the People's Republic of China have been increasing. Emissions from China in 2025 were 3.3 Pg C yr^{-1}, representing 30% of the global total.

In almost all global and regional carbon flux estimation systems, including CarbonTracker, fossil fuel CO_2 emissions are not optimized. Instead, these emissions are imposed and are not subject to revision by the inverse modeling framework. Global mass balance requires that any errors in fossil fuel emissions be compensated by opposing errors in land and ocean CO_2 exchange. Thus it is vital that fossil fuel CO_2

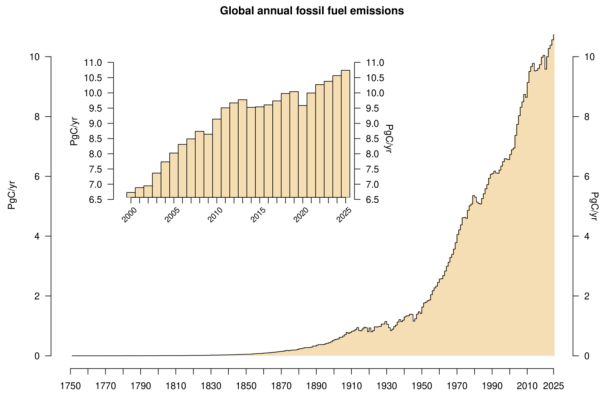


Figure 3.1: Time series of annual global fossil fuel emissions, in units of Pg C yr^{-1} (billion metric tons of C per year). Values from 1751 to 2014 are from Boden et al. (2017), and later values are extrapolated using consumption growth rate data of British Petroleum (2019). Inset figure shows the CT2026 period of analysis, 2000-2025.

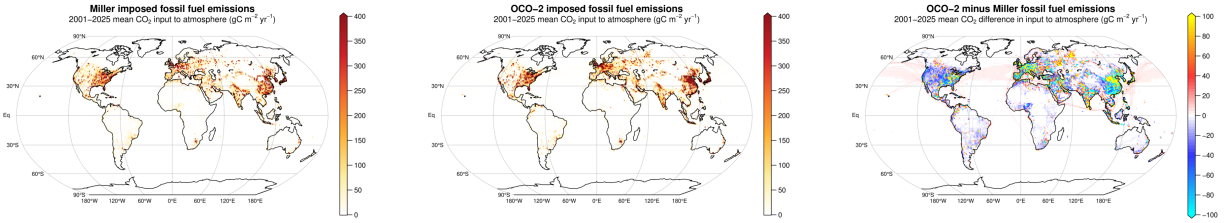
emissions are prescribed accurately, so that flux estimates for the land biosphere and oceans are robust. The fossil fuel emissions source data we use are available on an annually-integrated global and national basis. This aggregate information needs to be gridded before being incorporated into CarbonTracker. The major uncertainty in this process is distributing the national-annual emissions spatially across a nation and temporally into hourly contributions. In CT2026, we use two emissions products, called the "Miller" and "OCO-2" emissions datasets.

3.1.1 The "Miller" emissions dataset

Global and National Totals The Miller fossil fuel emission inventory is derived from independent global total and spatially-resolved inventories (Pera et al., 2026). Processing code for this product is available at `https://github.com/noaa-gml/Miller-Pera-FF`. Annual global total fossil fuel CO_2 emissions are based on the Appalachian Energy Center's "CDIAC at AppState" project (https://rieee.appstate.edu/projects-programs/cdiac/; Erb and Marland, 2026), which is an effort to update the original annual global and country fossil fuel-CO_2 emissions estimates from the DOE's Carbon Dioxide Information and Analysis Center (CDIAC) (Boden et al., 2017). The CDIAC at AppState emissions estimates used in CT2026 extend through 2022.

Spatial Distribution Miller fossil-fuel CO_2 fluxes are spatially distributed in two steps: First, the coarse-scale country totals through 2022 from CDIAC at AppState are mapped onto a $1^\circ \times 1^\circ$ grid according to the spatial patterns from the EDGAR 2025 GHG release (Crippa et al., 2025). The spatial pattern varies by year up until the end of the EDGAR product in 2024. After this, the 2024 spatial pattern is held constant. Note that while EDGAR provides annual emissions estimates at $1^\circ \times 1^\circ$ resolution, their totals do not agree with those from CDIAC at AppState. Thus, only the spatial patterns in EDGAR are used, and the total emissions are rescaled to CDIAC values. The CDIAC country-by-country totals sum to about 95% of the global total emissions; the remaining 5% is mapped to global shipping routes according to EDGAR, which we treat as a proxy for bunker fuel emissions.

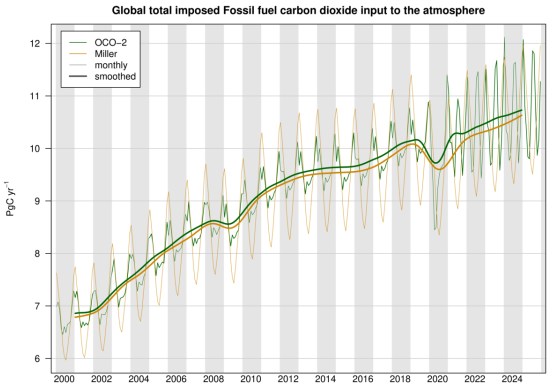
Temporal Distribution For North America between 30 and 60°N, the Miller system imposes a seasonal cycle derived from the first and second harmonics (Thoning et al., 1989) of the Blasing et al. (2004) analysis for the United States. The Blasing analysis has ˜10% higher emissions in winter than in summer. This scheme defines a fixed fraction of emissions for each month, so while the shape of



(a) Miller long-term average prior fossil-fuel CO_2 emissions

(b) OCO-2 long-term average prior fossil-fuel CO_2 emissions

(c) Miller minus OCO-2 long-term average difference in fossil-fuel CO_2 emissions.



(d) Miller minus OCO-2 long-term average difference in fossil-fuel CO2 emissions.

Figure 3.2: Imposed emissions of fossil-fuel CO_2 from the Miller and OCO-2 datasets (a) Miller fossil fuel long-term average CO_2 emissions in units of $g\,C\,m^{-2}\,yr^{-1}$. (b) As in panel a, but for the OCO-2 dataset. (c) Miller minus OCO-2 difference in long-term average fossil emissions. Warm (cool) colors indicate that Miller has a greater (smaller) source than OCO-2. (d) Global CO_2 fossil fuel emissions from the Miller (tan) and OCO-2 (green) priors, expressed in $Pg\,C\,yr^{-1}$. Thin lines are monthly averages, and thick lines represent data that have been smoothed with a 25-month Hanning window low-pass filter. Both datasets show similar trends of increasing emissions, and have similar interannual variability (notably an anomaly of decreased emissions during the COVID-19 pandemic). However, the Miller emissions generally have a greater annual cycle amplitude than the OCO-2 emissions.

the annual cycle is invariant, the amplitude of that cycle scales with the annual total emissions. For Eurasia, a set of seasonal emissions factors from EDGAR distributed by emissions sector is used to define fossil fuel seasonality. As in North America, this seasonality is imposed only from 30-60°N. The Eurasian seasonal amplitude is about 25%, significantly larger than that in North America, owing to the absence of a secondary summertime maximum due to air conditioning. See Figure 3.1 for the resulting time series of fossil fuel emissions. In order to avoid discontinuities in the fossil fuel emissions between consecutive years, a spline curve that conserves annual totals (Rasmussen, 1991) is fit to the annual emissions in each $1^\circ \times 1^\circ$ grid cell before the seasonal cycle is imposed.

Whereas early CarbonTracker releases used monthly-constant fossil fuel emissions, starting with CT2015 we introduced the use of temporal scaling factors to simulate day-of-week and diurnal variability for those emissions. These "Temporal Improvements for Modeling Emissions by Scaling" (TIMES) scaling factors, introduced by Nassar et al. (2013), are again applied to the Miller monthly emissions estimates for CT2026. The scaling factors consist of seven day-of-week global scaling factor maps, and 24 hourly global scaling factor maps to represent the diurnal cycle. For use in TM5,

the hourly scaling factors were aggregated to three-hourly factors to accommodate the time step of the model.

Extrapolation The full CDIAC at AppState dataset is only available through 2022 at time of running. A prior estimate of the fluxes through April of 2026 is required to optimize fluxes through the end of 2025, to accommodate our 12-week assimilation window.

For 2023 and 2024 the fractional increases of per-country, sectoral emissions are taken from The Energy Institute (formerly British Petroleum) Statistical Review of World Energy (Energy Institute, 2025) (for coal, oil, gas, and flaring) and the USGS NMIC Cement Mineral Commodity Summaries (for cement emissions). For example, to calculate 2023 coal emissions for France, the EI ratio of 2023 to 2022 coal consumption for France is used to scale 2022 CDIAC emissions, which are then used to rescale the EDGAR spatial patterns. Only the largest producing countries are included in the EI and USGS datasets, so regional or world-average ratios are used for extrapolation of other countries.

For 2025 and the beginning of 2026, global fractional increases are assumed (+2.5% per year for oil and gas, +1% per year for coal and flaring). The final months, February through April 2026, are further scaled by year-on-year ratios from the Carbon Monitor near-real-time estimates (Liu et al., 2020). No fuel-type data is available from Carbon Monitor, so the same year-on-year fractional changes are applied to all fuel types.

3.1.2 The "OCO-2" emissions dataset

The OCO-2 fossil fuel emissions (Basu and Nassar, 2026) were constructed primarily to be used for the OCO-2 Model Intercomparison Project (MIP). We use version 2026.2 of this OCO-2 MIP product. These data are available at https://zenodo.org/records/19744619.

For 2000–2023, these emission are based on ODIAC 2024 (Tomohiro Oda, 2015), which in turn uses BP's energy use statistics for 2022 and 2023. ODIAC monthly emissions have been disaggregated to hourly using the TIMES emission factors for day of week and time of day (Nassar et al., 2013). For 2024 onwards, ODIAC's 2023 emissions were scaled by the ratio of that month to 2023 emissions reported by Carbon Monitor, downloaded on April 24, 2026 from https://carbonmonitor.org/. ODIAC does not have sectoral decomposition to the degree provided by Carbon Monitor, so total ODIAC emissions for each region were scaled by the total emission change between 2023 and each extended year reported by Carbon Monitor. This means that power, ground transport, etc., have not been separately scaled.

Carbon Monitor data are daily, but ODIAC emissions are monthly. Therefore, Carbon Monitor data have been aggregated to monthly totals before deriving scaling factors between 2023 and the extended years. Carbon Monitor reports international aviation emissions by country of origin, while ODIAC reports aviation emissions on a grid. Since there is no way to derive the points of emission for Carbon Monitor aviation emissions, all Carbon Monitor international aviation was aggregated to create a single number for each month, then that number was used to scale ODIAC's bunker fuel for each month in 2024 and 2025.

CarbonTracker also requires fossil emissions for the first four months of 2026. However, the OCO-2 v2026.2 product only gives estimates through the end of February 2026. To extrapolate to the needed March and April days in 2026, we computed the 2026–2025 difference for the available days (in January and February). The average of this difference was computed, and then added to the corresponding March and April days of 2025 to make emissions in 2026.

3.1.3 Fossil fuel emissions datasets comparisons

Both the Miller and OCO-2 fossil fuel emissions datasets are characterized by an overall increase over the CarbonTracker analysis period (Figure 3.2d). In 2000, global emissions were about 6.7 Pg C yr^{-1}, and by 2025 they have risen to almost 11 Pg C yr^{-1}. Over this period, the OCO-2 emissions are systematically larger by about 0.1 Pg C yr^{-1}. Both emissions datasets manifest a similar dip in 2020 owing to reduced industrial and general transportation activity during the COVID-19 pandemic.

The annual cycle of fossil emissions is strikingly different between these two datasets. In general, the Miller emissions have a regular seasonal cycle throughout the record with a peak-to-peak amplitude larger than the OCO-2 ones. However, in the years after 2020, the ODIAC emissions on which the OCO-2 emissions are based have not only a greater seasonality, but also a marked mid-year peak in emissions. This difference in ODIAC seasonality exists in the original data as distributed at 1° x 1°, and is an artifact of how different sectoral emissions are extrapolated after the end of the Erb and Marland (2026) dataset (Tom Oda, private communication, 2026).

While the spatial distribution of long-term emissions as shown in Figures 3.2a and 3.2b is quite similar with emissions concentrated in industrialized countries and the most populous areas, the difference map (Figure 3.2c) shows that the OCO-2 dataset generally predicts more emissions over North America, Europe, and China compared to the Miller emissions. These differences in emissions propagate through the CarbonTracker inversion system to cause differences in inferred land and ocean fluxes, although it is difficult to identify specific effects of the fossil emissions differences.

3.1.4 Uncertainties

Marland (2008) attached an uncertainty of about 5% (95% confidence interval; approximately 2-σ) to the global total fossil fuel source. Estimates by Andres et al. (2014) put a larger uncertainty of 8.4% (2-σ) on the CDIAC global total. Uncertainties for individual regions of the world, and for sub-annual time periods are likely to be larger. Additional uncertainties are introduced when the emissions are distributed in space and time. In the Miller dataset, the overall Eurasian seasonality is based on scaling factors derived only from Western Europe and thus highly uncertain, but most likely a better representation than assuming no emission seasonality at all. Similarly, the use of the CDIAC monthly emission dataset for modeling seasonality introduces additional uncertainty in ODIAC. The additional uncertainty for the global total in the monthly CDIAC emission, which is solely due to the method for estimating seasonality, is reported as 6.4% (Andres et al., 2011). As mentioned earlier, fossil fuel emissions are not optimized in the current CarbonTracker system, similar to nearly all carbon data analysis systems. Spatial and temporal atmospheric CO_2 gradients arise from terrestrial biosphere and fossil-fuel sources. These gradients, which are interpreted by CarbonTracker, are difficult to attribute to one or the other cause. This is because atmospheric sampling sites have historically been established in locations remote from biospheric and anthropogenic sources, especially in the temperate Northern Hemisphere. Given that surface CO_2 flux due to biospheric activity and oceanic exchange is much more uncertain compared to fossil fuel emissions, CarbonTracker, like most current carbon dioxide data assimilation systems, does not attempt to optimize fossil fuel emissions. That is, the contribution of CO_2 from fossil fuel burning to observed CO_2 mole fractions is considered known. As detailed above, however, in CarbonTracker an effort is made to account for some aspects of fossil fuel uncertainty by using two different fossil fuel estimates. From a technical point of view, extra land biosphere prior flux uncertainty is included in the system to represent the random errors in fossil fuel emissions. Eventually, fossil fuel emissions could be optimized within CarbonTracker, especially with the addition of $^{14}CO_2$ observations as constraints (Basu et al., 2016, 2020).

3.2 Terrestrial biosphere module

The biospheric component of the terrestrial carbon cycle consists of all the carbon stored in 'biomass' around us. This includes trees, shrubs, grasses, carbon within soils, dead wood, and leaf litter. Such reservoirs of carbon exchange CO_2 with the atmosphere. This exchange includes plants taking up CO_2 during their growing season, via photosynthesis. Most of this carbon is released back to the atmosphere throughout the year through the process of respiration. Respiration is the conversion of organic carbon back to inorganic CO_2 by plants, animals, and microbes. It includes both the decay of dead wood and litter (heterotrophic respiration), and the metabolic respiration of living plants (autotrophic respiration).

Plants can also return carbon to the atmosphere when they burn, as described in Section 3.3. Even though the yearly sum of uptake and release of carbon amounts to a relatively small number, a few petagrams (one Pg=10^{15} g of carbon per year), the flow of carbon each way is as large as 120 Pg C each year. These flows are each affected differently by changes in temperature, water availability, and other factors.

The net result of these land biosphere flows needs to be monitored because of its significant impact on atmospheric CO_2 levels. Due to the relative complexity of the terrestrial carbon cycle, we need a good physical description—a model—of these flows of carbon. This need derives in part from the fact that atmospheric measurements of CO_2 see only the relatively small net sum of the much larger two-way streams, or gross fluxes. Information on what the biospheric fluxes are doing in each season, and in every location on Earth, is derived from specialized land biosphere models, and fed into our system as a first guess, to be refined by our assimilation procedure.

In CT2026, we use the MiCASA and SiB4 terrestrial biosphere models to simulate net ecosystem exchange (NEE), which is the sum of photosynthetic uptake of atmospheric CO_2, and the return of CO_2 to the atmosphere through respiration.

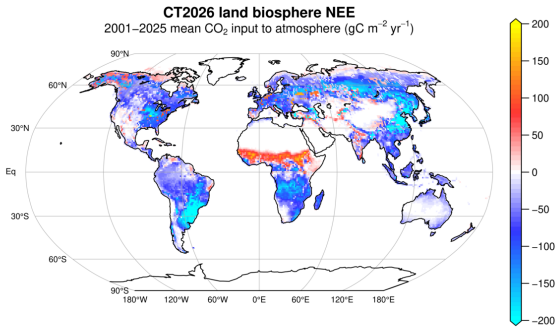
In addition to simulating NEE, each of these models is also coupled to a fire model. Fire models calculate the amount of carbon released to the atmosphere via wildfire and other non-fossil-fuel combustion. As it turns out, both systems use fire models based on the Global Fire Emissions Database (van der Werf et al., 2004, and Section 3.3). As described below, MiCASA uses a modified version of GFED3, and SiB4 imposes its fire emissions directly from the GFED4.1s product.

3.2.1 MiCASA model

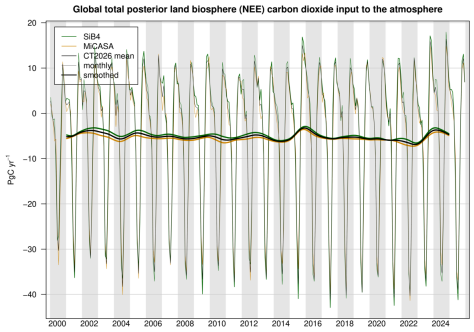
First-guess terrestrial biosphere fluxes for CT2026 are provided by the NASA Más informada Carnegie-Ames-Stanford-Approach (MiCASA) v1 model. This model is based on the CASA model introduced by Potter et al. (1993), and provides monthly net primary production (fixation of atmospheric CO_2 via photosynthesis less autotrophic respiration) and heterotrophic respiration at 0.1° x 0.1° lateral resolution. For use in CarbonTracker, these data are spatially averaged to 1° x 1° and temporally downscaled as described in Section 3.2.2. The MiCASA product is documented at https://acdisc.gsfc.nasa.gov/data/CMS/MICASA_FLUX_D.1/doc/MiCASA_README.pdf and is available for download at https://portal.nccs.nasa.gov/datashare/gmao/geos_carb/MiCASA/v1/.

We use the standard v1 product when it exists and the near-real-time vNRT stream for the period that v1 has not yet reached. MiCASA begins in 2001. For 2000 we use the climatology of NPP and heterotrophic respiration computed from the 2001–2025 record. This is also how we handle MiCASA fire emissions in 2000 (Section 3.3.2).

Note that the NASA MiCASA team produces temporally-downscaled GPP, heterotrophic respiration, and



(a) Long-term mean posterior CT2026 NEE fluxes in g C m^{-2} yr^{-1}.



(b) Global land uptake by the two inversions

Figure 3.3: Posterior NEE from CT2026. (a) Long-term mean posterior CT2026 NEE fluxes in g C m^{-2} yr^{-1}. Negative fluxes (cool colors) represent CO_2 uptake by the land, whereas positive fluxes (warm colors) indicate regions in which the land is a net source of CO_2 to the atmosphere. NEE does not include fire emissions. (b) Global land uptake by the two inversions using MiCASA (tan), SiB4 (green) and the final CT2026 average (black). Monthly values are in thin lines, and fluxes smoothed using a 25-month Hanning low-pass filter are in thick lines.

fires with 3-hourly resolution. This is done using MERRA-2 meteorology and a scheme similar to Olsen and Randerson (2004). We do not use this downscaled product, in part because the MERRA-2 meteorology is different from the ECMWF meteorology, and in part because the spatial resolution of the MERRA-2 meteorology is different from our 1° × 1° flux grid. Please see Section 3.2.2 below about our temporal downscaling scheme.

The MiCASA team provides an atmospheric correction term computed following the methods described in Weir et al. (2021). This is intended to be added to the native MiCASA flux components to create a terrestrial prior with a sink much closer to that required by atmospheric growth rates. One condition of Bayesian methods is that the prior be unbiased, and this correction term is intended to meet this condition for the MiCASA product. CarbonTracker's state vector is in scaling factor space instead of flux space, and the condition of having unbiased priors is primarily met by using projected scaling factors. CarbonTracker was in fact designed to correct terrestrial biosphere models with an insufficient land sink. As a result, we do not use this atmospheric correction term.

CASA models directly simulate monthly-mean Net Primary Production (NPP) and heteotrophic respiration (R_H) for each terrestrial grid cell being simulated. NPP is the difference in photosynthetic carbon uptake (Gross Primary Production, GPP) and the carbon release by the same plants due to "maintenance respiration", which is also called autotrophic respiration, R_A. The carbon uptake represented by NPP and carbon release represented by R_H can be differenced to provide Net Ecosystem Exchange (NEE) of CO_2. Throughout this discussion, we use the convention that fluxes carry algebraic signs and we adopt the "atmospheric perspective" for those signs. Thus carbon uptake by the terrestrial biosphere is a negative flux to the atmosphere, and release of CO_2 back to the atmosphere is a positive flux. This means that we represent all respiration fluxes as positive and GPP as negative, so NEE = NPP + R_H. This stands in contrast to convention in the terrestrial carbon community, where all fluxes are generally non-negative.

3.2.2 Temporal downscaling

Use of monthly-mean terrestrial fluxes to simulate atmospheric CO_2 is not sufficient to resolve the variability observed at measurement sites. Instead, higher-frequency variations, including the diurnal cycle and effects of passing weather systems must be imposed on the MiCASA monthly fluxes. Following the logic laid out by Olsen and Randerson (2004), we transform the MiCASA-supplied monthly-mean NPP and R_H fluxes into GPP and total ecosystem respiration, $R_E = R_A + R_H$.

To estimate sub-monthly variations, including diurnal and synoptic variability, the Olsen and Randerson (2004) strategy is to model GPP as a linear function of incoming surface solar radiation and total ecosystem respiration as a function of near-surface temperature.

The fundamental assumption needed to apply this scheme is that we can resolve MiCASA-simulated NPP into GPP and R_A. We apply the assumption that GPP is twice NPP, which further implies that R_A is the same size as NPP, but of opposite sign:

$$\text{GPP} = 2 * \text{NPP}, \tag{3.1}$$

$$\text{NPP} = \text{GPP} + R_A, \tag{3.2}$$

and

$$R_A = -1 * \text{NPP}. \tag{3.3}$$

This fixes the ratio of GPP to autotrophic respiration at 2:1 globally. This is the largest structural assumption in the current and Olsen and Randerson (2004) schemes.

We use meteorological fields from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis to supply fields of surface temperature and shortwave radiation. Fluxes are generated at hourly resolution using a simple temperature Q_{10} relationship for respiration, assuming a global Q_{10} value of 1.5 applied to 2-meter air temperature, and a linear scaling of photosynthesis with solar radiation. The procedure is very similar, but not identical to the procedure in Olsen and Randerson (2004). Note that the introduction of hourly variability conserves the monthly mean NEE from MiCASA. Instantaneous NEE for each hour is created as:

$$\text{NEE}(t) = \text{GPP}(t) + R_E(t), \tag{3.4}$$

where

$$\text{GPP}(t) = \text{GPP}_{\text{mean}}(I(t)/I_{\text{mean}}) \tag{3.5}$$

$$R_E(t) = R_{E,\text{mean}}(Q_{10}(t)/Q_{10,\text{mean}}), \tag{3.6}$$

and Q_{10} is computed as

$$Q_{10}(t) = 1.5^{(T_{2m}(t)-273.15)/10.0}, \tag{3.7}$$

where T_{2m} is temperature at 2 meters above the land surface in Kelvin, I is surface incoming solar radiation, t is time in hourly intervals, and x_{mean} represents the monthly mean of quantity x, including the monthly-mean fluxes derived from MiCASA. Whenever incoming shortwave radiation is zero (for example, during polar night), GPP is forced to zero, and the uptake removed from those hours is redistributed onto the cell's sunlit hours so that the monthly mean is preserved exactly.

Smooth month-to-month variations

While the scheme outlined above imposes realistic diurnal- and synoptic-scale variations on monthly-mean GPP and R_E, it still allows for abrupt changes from one month to the next. For CT2026, we add a further processing step designed to remove such unrealistic step changes. We fit smooth curves to the timeseries of monthly GPP and R_E using the monotone piecewise cubic Hermite interpolation (Fritsch and Carlson, 1980). These fits are piecewise quadratic curves that reproduce every monthly mean exactly and are continuous across month boundaries. We use a similar scheme to smooth over year-to-year step changes in fossil fuel emissions. The final smoothed GPP is

$$\text{GPP}_F(t) = \text{GPP}(t) - \text{GPP}_{\text{mean}} + \text{GPP}_{\text{FIT}}(t), \tag{3.8}$$

and the final smoothed ecosystem respiration is

$$R_{E,F}(t) = R_E(t) - R_{E,\text{mean}} + R_{E,\text{FIT}}(t). \tag{3.9}$$

Together, these form the terrestrial NEE imposed as a first-guess flux in CT2026:

$$\text{NEE}_F(t) = \text{GPP}_F(t) + R_{E,F}(t). \tag{3.10}$$

3.2.3 SiB4

The Simple Biosphere Model (SiB4; Haynes et al., 2019a,b) is a mechanistic and process-based model that simulates land-atmosphere exchanges of energy, momentum and moisture, as well as the terrestrial carbon cycle over heterogeneous vegetation. SiB4 combines elements from a prognostic phenology model (Stöckli et al., 2011), a crop model Lokupitiya et al. (2016), and a terrestrial carbon pool model (SiB-CASA; Schaefer et al., 2008). To capture vegetation-specific phenology and biological processes, SiB4 uses 19 plant functional types (PFTs): eight tree PFTs (four types further sub-divided by existence in tropical, temperate, and boreal regions), two shrublands, four grasslands, four crops, and a desert/bare ground classification.

SiB4 was run using ERA5 meteorological drivers, which is the same meteorology as our TM5 transport model. Fluxes were natively produced at scales chosen specifically for CarbonTracker. Spatially, they are produced at 1° x 1° and they have hourly resolution from 2000 through the first few months of 2026.

To simulate land cover heterogeneity, SiB4 uses tiles to allow multiple PFTs in a single grid cell. The magnitude of carbon taken up during photosynthesis depends directly on environmental factors and the fraction of absorbed photosynthetically active radiation. Photosynthesis is coupled to the energy budget,

transpiration, and Bowen ratio following Collatz et al. (1991, 1992). Uptake from photosynthesis is either respired for maintenance or allocated to a live carbon pool. Once daily, carbon from live carbon pools is distributed to dead carbon pools following a cascading-pool setup. Heterotrophic respiration resulting from the dead carbon pools is modulated by temperature, moisture, and turnover time.

SiB4 has a relatively aggressive net land sink. Its prior 2001–2025 mean global land sink is almost 4Pg C yr^{-1}. It is understood that this SiB4 land sink is mechanistically dominated by CO_2 fertilization in the photosynthesis scheme. The magnitude of this CO_2 fertilization effect can be estimated by comparing two runs of SiB4, one with a realistically-varying atmospheric CO_2 boundary condition and one with a constant CO_2. That comparison, an approximation of the CO_2 fertilization effect in SiB4, yields a 2000–2019 mean difference of 3 Pg C yr^{-1}, and a significant trend of about 0.7 Pg C yr^{-1} per decade.

Fire in SiB4

SiB4 imposes GFED4.1s fire emissions by removing burned biomass from its carbon pools. It dynamically allocates burning across its carbon pools following a pre-determined, rudimentary sequential burning of PFTs and carbon pools. Within a given 0.5-degree grid cell, any grasslands that are present are burned first. The remaining PFTs are burned in descending order of total above-ground biomass. Burning continues until the total grid cell emissions equal the GFED4.1s total for that given grid box and time step. In this way, global and regional fire totals from SiB4/GFED4.1s are equivalent to those of GFED4.1s, and fire causes a dynamic reduction and restructuring of carbon pools.

TODO need info on diurnal cycle in SiB4/GFED4.1s fires (Mu et al., 2011). Waiting on response from Sourish on code he wrote for Aleya.

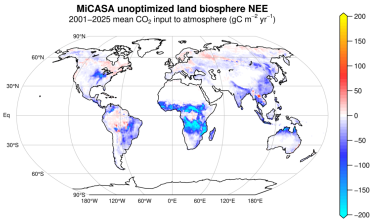
3.2.4 Comparison of terrestrial biosphere priors

The record of atmospheric CO_2 calls for a deeper terrestrial biosphere sink than that generally simulated by terrestrial biosphere models. Inverse models generally simulate a larger annual cycle of terrestrial biosphere fluxes, and in particular a deeper boreal summer uptake of carbon dioxide, in the posterior optimized fluxes compared to the prior models. We call upon the atmospheric CO_2 observations to make this change, and in order to handle these prior model differences the ensemble Kalman filter's prior covariance model has to be appropriately tuned. In short, this prior uncertainty needs to comfortably span differences among the terrestrial biosphere priors, the fossil fuel emissions estimates, and adjustments to fluxes required to bring model predictions into agreement with observations. CT2026 prior covariances have been adjusted compared to prior releases, and details on this adjustment can be found in Section 6.

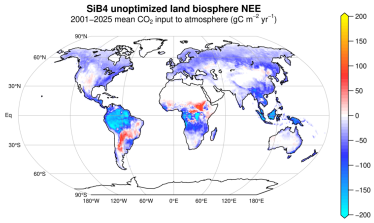
Due to the inclusion of fires, inter-annual variability in weather and NDVI, and increasing atmospheric CO_2 levels, NEE fluxes for land regions start with an uptake flux even before optimizing the fluxes. Globally-integrated, this first-guess flux is about -3.2 Pg C yr^{-1}, which is increased in magnitude to about -5.3 Pg C yr^{-1} by the optimization process. Note that negative fluxes represent a sink of atmospheric CO_2.

3.3 Fire module

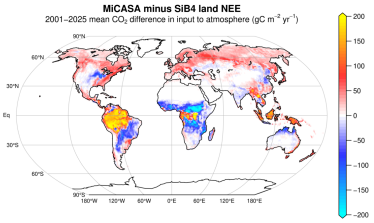
Vegetation fires are an important part of the carbon cycle and have been so for many millennia. Even before human civilization began to use fires to clear land for agricultural purposes, most ecosystems were subject



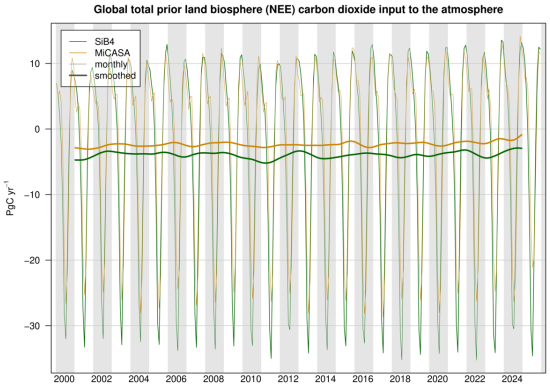
(a) MiCASA long-term average prior net ecosystem exchange.



(b) SiB4 long-term average prior net ecosystem exchange.



(c) MiCASA minus SiB4 long-term average difference in prior net ecosystem exchange.



(d) Global net ecosystem exchange by the MiCASA (tan) and SiB4 (green) priors, expressed in Pg C yr^{-1}. Thin lines are monthly-average fluxes, and thick lines have been smoothed with a 25-month Hanning window low-pass filter.

Figure 3.4: Prior net ecosystem exchange (NEE) of CO_2 from the MiCASA and SiB4 terrestrial biosphere models. NEE includes uptake by photosynthesis and release by respiration, but does not include fire sources. (a) MiCASA long-term average prior NEE in units of g C m^{-2} yr^{-1}. Warm colors indicate positive flux, representing release of CO_2 to the atmosphere. Cool colors show places where the land biosphere (excluding fires) is a sink of atmospheric CO_2. (b) As in panel a, but for the SiB4 prior. (c) MiCASA minus SiB4 difference in long-term average NEE. Warm colors indicate either that MiCASA has a greater land source than SiB4, or a smaller sink. Cool colors indicate the opposite. (d) Global NEE by the MiCASA (tan) and SiB4 (green) priors, expressed in Pg C yr^{-1}.

to natural wildfires that would rejuvenate old forests and bring important minerals to the soils. When fires consume part of the landscape in either controlled or natural burning, carbon dioxide (among many other gases and aerosols) is released in large quantities. Each year, vegetation fires emit around 2-3 Pg C as CO_2 into the atmosphere, mostly in the tropics.

Currently, a large fraction of wildfire is started by humans. This is mostly intentional to clear land for agriculture, or to re-fertilize soils before a new growing season. This important component of the carbon cycle is monitored mostly from space, while sophisticated 'biomass burning' models are used to estimate the amount of CO_2 emitted by each fire. Such estimates are then used in CarbonTracker to prescribe the emissions. These emissions are not modified in the optimization (inverse modeling) process.

3.3.1 GFED fire emissions model

The Global Fire Emissions Database (GFED) was introduced by van der Werf et al. (2004). GFED is driven by satellite observations of fire occurrence, and determines emissions based on carbon pool sizes where fires occur, and indicators of combustion completeness. These carbon pools are provided by a terrestrial biosphere model coupled to GFED. Burned mass is converted to CO_2 emissions using pre-determined emissions factors.

3.3.2 The MiCASA GFED v3 variant

MiCASA v1 simulates biomass burning using a successor model to the CASA-GFED3 model of van der Werf et al. (2004). The NASA MiCASA team produces emissions of wildfire and wood fuel combustion using Moderate Resolution Imaging Spectroradiometer (MODIS) Normalized Difference Vegetation Index (NDVI) computed from reflectances in product MCD43A4.061, burned area from product MCD64A1.061, and carbon fuel stock in various biomass pools estimated from the CASA biogeochemical model. Fire and fuel emissions are available on a daily basis from 2001-2025. For 2000 we apply the climatology of MiCASA fire emissions, computed from its 2001-2025 mean.

In this GFED variant, burned area is based on MODIS satellite observations of fire counts. These, together with detailed vegetation cover information and a set of vegetation specific scaling factors, allow predictions of burned area when active fire counts from MODIS are available. The relationship between fire counts and burned area is derived, for the specific vegetation types, from a calibration subset of 500m resolution burned area from MODIS in the period 2001-2004.

Once burned area has been estimated globally, emissions of trace gases are calculated using carbon pool stocks from the CASA biosphere model. The seasonally changing vegetation and soil biomass stocks in the CASA model are combusted based on the burned area estimate, and converted to atmospheric trace gases using estimates of fuel loads, combustion completeness, and burning efficiency.

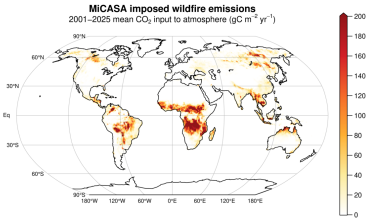
3.3.3 GFED4.1s fire emissions

The GFED4 product (Randerson et al., 2017) is an evolution of the Global Fire Emissions Database version 3 (GFED3) biogeochemical model (Randerson et al., 2012). It was later augmented with estimates of emissions from smaller fires not directly seen by the original satellite observations to form the GFED4.1s product.

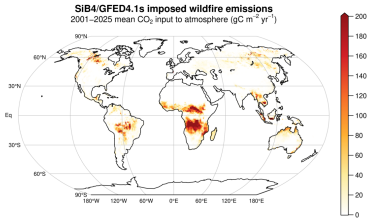
The GFED4.1s burned area estimate is documented in Giglio et al. (2013), and includes an experimental estimate of small fires. From 2001 onwards, this combines burned area maps from MODIS product MCD64A1 with active fire data from the Tropical Rainfall Measuring Mission (TRMM) Visible and Infrared Scanner (VIRS) and the Along-Track Scanning Radiometer (ATSR) family of sensors. Before 2001, ATSR and VIRS data alone were used. Burned area from small fires depends on active fire detection from MODIS thermal anomalies as described in Randerson et al. (2012). The burned area estimates were converted to emissions using the GFED3 biogeochemical model (Randerson et al., 2012).

3.3.4 Fire emissions comparison

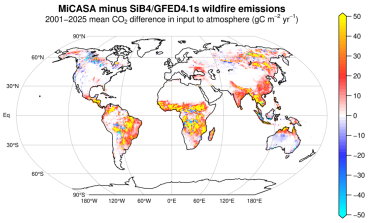
MiCASA is characterized by a significantly larger global fire emissions of about 3.3 Pg C yr^{-1} averaged over 2000-2025 compared to those of SiB4/GFED4.1s, at 2.1 Pg C yr^{-1} (Figure 3.5d). This is in part due



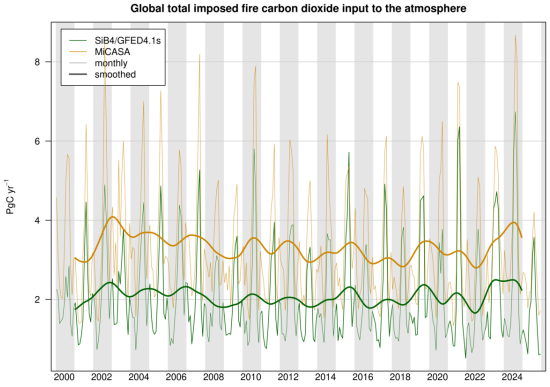
(a) MiCASA long-term average prior fire emissions



(b) SiB4/GFED4.1s long-term average prior fire emissions



(c) MiCASA minus SiB4/GFED4.1s long-term average difference in fire emissions.



(d) Global emissions of CO_2 from biomass burning prescribed by the MiCASA (tan) and SiB4/GFED4.1s (green) priors, expressed in Pg C yr^{-1}.

Figure 3.5: Imposed emissions of fire CO_2 from the MiCASA and SiB4/GFED4.1s datasets (a) MiCASA long-term average fire emissions in units of g C m^{-2} yr^{-1}. (b) As in panel a, but for the SiB4/GFED4.1s dataset. (c) MiCASA minus SiB4/GFED4.1s difference in long-term average fire emissions. Warm (cool) colors indicate that MiCASA has a greater (smaller) source than SiB4/GFED4.1s. (d) Global biomass-burning emissions from the MiCASA (tan) and SiB4/GFED4.1s (green) priors, expressed in Pg C yr^{-1}. Thin lines are monthly averages, and thick lines represent data that have been smoothed with a 25-month Hanning window low-pass filter. MICASA simulates about 1.25 Pg C yr^{-1} more fire emissions than SiB4/GFED4.1s. However, they both manifest similar interannual variability.

to the inclusion of an explicit biofuel-burning component in MiCASA, although this only explains about 0.4 Pg C yr^{-1} of the difference. The MiCASA also simulates larger non-biofuel fire emissions of around 2.9 Pg C yr^{-1}.

The relatively new GFED5 product (van der Werf et al., 2025) computes emissions based on observations of burned area. This approach yields emissions of CO_2 are significantly greater than those of GFED4.1s. A version of SiB4 using burned area is under development, with the constraint that total burned area in both models is conserved. The fire emissions in this new version of SiB4 will come from the product of burned area, pool biomass, and combustion completeness and mortality factors that are modulated by environmental conditions such as aridity.

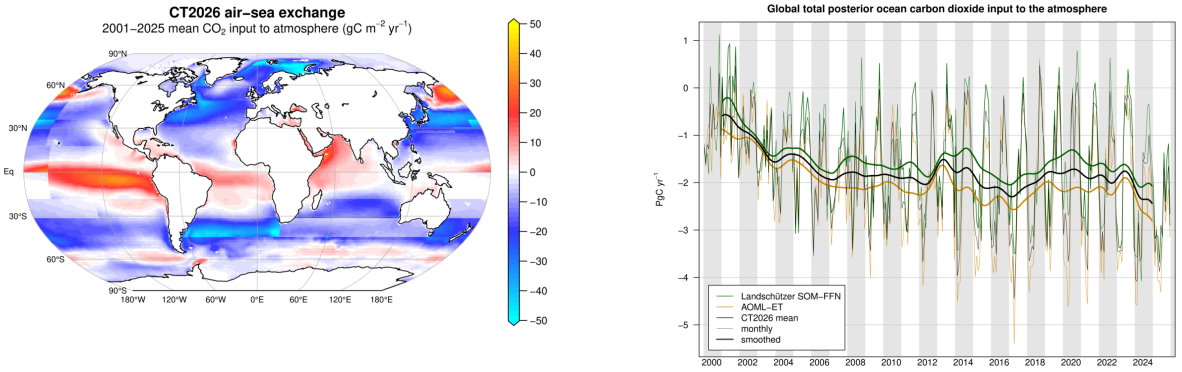
The relatively large discrepancy in fire emissions between the fire products used in CT2026 increases overall uncertainty in optimized flux results, especially for non-fire land emissions (NEE). The inversion using MiCASA priors estimates a posterior global land NEE of about -5.7 Pg C yr^{-1}, whereas the inversion using SiB4/GFED4.1s priors finds about -5 Pg C yr^{-1}. The remaining difference ends up being ascribed to global ocean fluxes.

3.4 Oceans module

The oceans play an important role in the Earth's carbon cycle. They are the largest long-term sink for carbon and have an enormous capacity to store and redistribute CO_2 within the Earth system. Oceanographers estimate that about 48% of the CO_2 from fossil fuel burning has been absorbed by the ocean (Sabine et al., 2004). The global ocean CO_2 sink was 2.9 ± 0.4 Pg C yr^{-1} during the decade 2014-2023, representing 26% of total CO_2 fossil fuel and land use emissions (Friedlingstein et al., 2024). The dissolution of CO_2 in seawater shifts the balance of the ocean carbonate equilibrium towards a more acidic state with a lower pH. This effect is already measurable (Caldeira and Wickett, 2003), and is expected to become an acute challenge to shell-forming organisms over the coming decades and centuries. Although the oceans as a whole have been a reliable sink of carbon over the past decades, CO_2 is also be released from regions of the ocean depending on local temperatures, biological activity, wind speeds, and ocean circulation. These processes are all considered in CarbonTracker, since they can have significant effects on the ocean sink. Improved estimates of the air-sea exchange of carbon in turn help us to understand variability of both the atmospheric burden of CO_2 and terrestrial carbon exchange.

Carbontracker CT2026 finds that the ocean sink begins in 2000 at about about 1 Pg C yr^{-1}, but by 2025 has exceeded 2 Pg C yr^{-1} (cf. Figure 3.6b and the CT2026 website). This secular increase is modulated by significant interannual variability.

The spatial distribution of CT2026 air-sea exchange of CO_2 is strong extratropical uptake in both the northern and southern hemispheres. Meanwhile, in agreement with our understanding of the ocean carbon cycle, equatorial upwelling of deep CO_2-rich waters leads to an tropical source of carbon to the atmosphere (Figure 3.6a.) This map also reveals some of the internal workings of CarbonTracker, and its limitations. Region-to-region discontinuities are the result of so-called "flux dipoles", where available observational constraints are not sufficient to uniquely and independently constrain flux from certain regions. In such situations, neighboring regions' fluxes will play off against one another. This is indicative of CarbonTracker being able to better constrain their sum than each region individually. More well-behaved results can be seen by plotting the flux from larger ocean regions, as in this plot.



(a) **(b)**

Figure 3.6: Posterior CT2026 ocean fluxes. (a) Long-term mean posterior ocean fluxes in g C m^{-2} yr^{-1}. Negative fluxes (cool colors) represent CO_2 uptake by the ocean, whereas positive fluxes (warm colors) indicate regions in which the ocean is a net source of CO_2 to the atmosphere. Discontinuities between ocean regions are indicative of flux dipoles (see discussion). (b) Global ocean uptake by the two inversions using SOM-FFM (green), AOML-ET (tan) and the final CT2026 average (black). Monthly values are in thin lines, and fluxes smoothed using a 25-month Hanning low-pass filter are in thick lines.

3.4.1 Air-sea CO_2 exchange

Oceanic uptake of CO_2 in CarbonTracker is computed using differences in partial pressure of CO_2 between the atmosphere and the ocean surface. The global distribution of seawater partial pressure of CO_2, denoted pCO_2, is inferred from analysis of direct measurements of that quantity, followed by an interpolation procedure. The resulting global air-sea partial pressure differences are combined with a gas transfer velocity computed from wind speeds in the atmospheric transport model to compute fluxes of carbon dioxide across the sea surface.

In these gas-exchange computations it is formally correct to consider the fugacity of CO_2 instead of its partial pressure (Dickson et al., 2008). Fugacity can be thought of as an effective partial pressure meant to correct for slightly non-ideal gas behavior of CO_2. The difference between fugacity and partial pressure is very small, about 0.3%, and we neglect this correction in our computations.

In the following sections we first describe the SOM-FFN pCO_2 and AOML-ET prior models. We then describe the air-sea gas transfer velocity parameterization, compare the two priors, and discuss details of the inversion methodology specific to oceanic exchange of CO_2.

3.4.2 SOM-FFN model

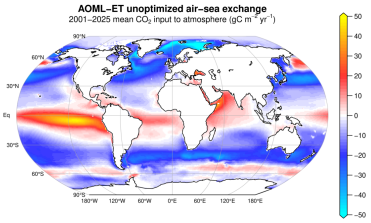
One estimate of surface ocean pCO_2 for CT2026 is provided by the SOM-FFN product of Landschützer et al. (2013). This is a two-step machine-learning system assimilating gridded monthly seawater pCO_2 measurements from the Surface Ocean CO_2 Atlas project (SOCAT; Bakker et al., 2016). The method first clusters the global ocean into biogeochemical provinces using a self-organized maps method (Kohonen, 1987). In a second step, it reconstructs the nonlinear relationship between driver variables and pCO_2 observations using a feed-forward neural network. In this way, available measurements are mapped to a globally-complete 1° x 1° monthly ocean grid (Landschützer et al., 2016).

CT2026 uses the SOM-FFN version produced for the 2025 Global Carbon Budget (Friedlingstein et al., 2026) through the end of 2024, and an extrapolation thereafter. All SOM-FFN versions can be downloaded at NCEI (Jersild et al., 2017).

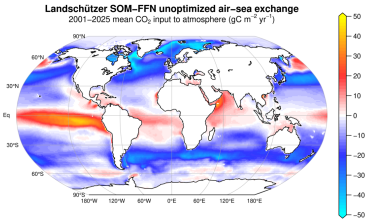
3.4.3 AOML-Extra Trees model

A second estimate of surface ocean pCO_2 for CT2026 is provided by the NOAA Atlantic Oceanographic and Meteorological Laboratory - Extremely Randomized Trees (AOML-ET; Wanninkhof et al., 2025) system. This is a machine-learning system that assimilates gridded monthly seawater pCO_2 measurements from the Surface Ocean CO_2 Atlas project (SOCAT; Bakker et al., 2016). These gridded measurements are mapped to a globally-complete 1° x 1° monthly ocean grid using a supervised regression method (Geurts et al., 2006) relying on implicit relationships between pCO_2 and the chosen regressors of time, location, sea surface temperature, sea surface salinity, chlorophyll, and mixed-layer depth. It provides a spatially and temporally resolved representation of the partial pressure of CO_2 in surface waters. This product accounts for variability in ocean surface properties, including temperature and biological activity, which influence the solubility and concentration of CO_2 in seawater.

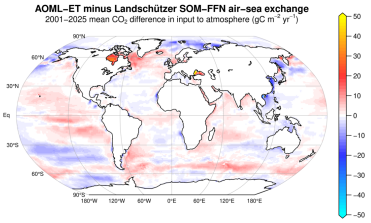
CT2026 uses AOML-ET version v20240425 through the end of 2023, and v20250617 thereafter. AOML-ET products are published at NCEI (Wanninkhof et al., 2024).



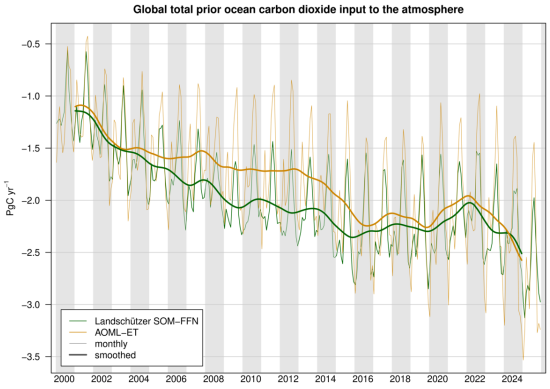
(a) AOML-ET long-term average prior air-sea CO_2 exchange.



(b) SOM-FFN long-term average prior air-sea CO_2 exchange



(c) AOML-ET minus SOM-FFN long-term average difference in prior air-sea CO_2 exchange.



(d) Global CO_2 uptake by the ocean by the SOM-FFN (green) and AOML-ET (tan) priors, expressed in Pg C yr^{-1}.

Figure 3.7: Prior air-sea exchange of CO_2 from simulations using the AOML-ET and SOM-FFN ocean pCO_2 fields, observed atmospheric CO_2 from the NOAA marine boundary layer surface, and a gas transfer scheme described in Section. 3.4.4. (a) AOML-ET long-term average prior air-sea CO_2 exchange in units of g C m^{-2} yr^{-1}. Warm colors indicate positive flux, representing release of CO_2 to the atmosphere. Cool colors show places where the ocean is a sink of atmospheric CO_2. (b) As in panel a, but for the SOM-FFM prior. (c) AOML-ET minus SOM-FFM difference in long-term average air-sea exchange. Warm colors indicate either that AOML-ET has a greater ocean source than SOM-FFN, or a smaller sink. Cool colors indicate the opposite. (d) Global CO_2 uptake by the ocean by the SOM-FFN (green) and AOML-ET (tan) priors, expressed in Pg C yr^{-1}. Both priors show similar trends of increasing oceanic uptake, and manifest similar interannual variability. However, the AOML-ET prior has a greater annual cycle than the SOM-FFN prior, and generally predicts a smaller ocean sink.

3.4.4 Gas-transfer velocity and ocean surface properties

Both priors use CO_2 solubilities and Schmidt numbers computed from World Ocean Atlas 2009 (WOA09) climatological fields of sea surface temperature and sea surface salinity fields (Levitus et al., 2010). Gas transfer velocity in CarbonTracker is parameterized as a quadratic function of wind speed following Wanninkhof (2014). Gas exchange is computed every timestep using wind speeds from the ERA5 reanalysis as represented by the atmospheric transport model.

Air-sea transfer is inhibited by the presence of sea ice, and for this work fluxes are scaled by the daily sea ice fraction in each gridbox provided by the ERA5 data.

3.4.5 Specifics of the inversion methodology related to air-sea CO_2 fluxes

The first-guess fluxes described here are subject to scaling during the CarbonTracker optimization process, in which atmospheric CO_2 mole fraction observations are combined with transport simulated by the atmospheric model to infer flux signals. Prior air-sea fluxes are adjusted within each of the 30 ocean inversion regions. In this process, signals of terrestrial flux in atmospheric CO_2 distribution can be erroneously interpreted as being caused by oceanic fluxes. This flux "aliasing" or "leakage" is evident in some regions as a change in the shape of the seasonal cycle of air-sea flux.

Uncertainty on the prior model is specified as uncertainties on scaling factors multiplying net CO_2 flux in each of the 30 ocean inversion regions. In the absence of a suitable uncertainty estimate from the AOML-ET prior model, we apply an uncertainty computed from the ocean interior inversion of Jacobson et al. (2007). This choice is preliminary and will be revised in the future.

3.4.6 Prior ocean emissions comparison

The overall pattern of average air-sea gas exchange as shown in Figures 3.7a and 3.7b is consistent with our understanding of the ocean's biogeochemical cycling of CO_2. While the global ocean is characterized by overall uptake of CO_2 in the extratropics, the equatorial zone is believed to be a source of CO_2 to the atmosphere (Takahashi et al., 2002). Differences between the two priors in spatial distribution of long-term uptake are generally less than about $\pm$ 10g C m^{-2} yr^{-1} (Figure 3.7c). The generally-red colors in this map indicate that over most of the ocean, AOML-ET predicts a smaller ocean sink compared to the SOM-FFN model.

The overall ocean sink is increasing in response to rising atmospheric CO_2 concentrations (Figure 3.7d). The SOM-FFN model consistently predicts a slightly larger ocean sink than AOML-ET (-2 Pg C yr^{-1} compared to AOML-ET's long-term average flux of -1.9 Pg C yr^{-1}). While we would hope that the inversion process would help us to converge on a final result for this ocean uptake, we find that the two inversions actually end up with more discrepancy on the long-term mean ocean sink size (-2 for AOML-ET and -1.6 Pg C yr^{-1} for SOM-FFN in the 2001-2025 mean global posterior flux). This is due to differences in the land priors between the two inversions (see Section 3.3.4), and the inability of available measurements and our system to uniquely quantify ocean uptake.

Chapter 4

Atmospheric transport

The link between observations of CO_2 in the atmosphere and the exchange of CO_2 at the Earth's surface is transport in the atmosphere: storm systems, cloud complexes, and weather of all sorts cause winds that transport CO_2 around the world. As a result, local surface CO_2 exchange events like fires, forest growth, and ocean upwelling can have impacts at remote locations. To simulate the winds and the weather, CarbonTracker uses sophisticated numerical models that are driven by the daily weather forecasts from the specialized meteorological centers of the world. Since CO_2 does not decay or react in the lower atmosphere, the influence of emissions and uptake in locations such as North America and Europe are ultimately seen in our measurements even at the South Pole. Getting the transport of CO_2 just right is an enormous challenge, and costs us almost all of the computer resources for CarbonTracker. To represent the atmospheric transport, we use the Transport Model 5 (TM5). This is a community-supported model whose development is shared among many scientific groups with different areas of expertise. TM5 is used for many applications other than CarbonTracker, including forecasting air-quality, studying the dispersion of aerosols in the tropics, tracking biomass burning plumes, and predicting pollution levels that future generations might have to deal with.

4.1 TM5 offline tracer transport model

TM5 is an offline global chemical transport model with two-way nested grids. In this global model, regions for which high-resolution simulations are desired can be nested in the coarser global grid. The advantage to this approach is that transport simulations can be performed with a regional focus without the need for boundary conditions. Further, this approach allows measurements outside the "zoom" domain to constrain regional fluxes in the data assimilation, and ensures that regional estimates are consistent with global constraints. TM5 is based on a predecessor model TM3, with improvements in the advection scheme, vertical diffusion parameterization, and meteorological preprocessing of the wind fields (Krol et al., 2005).

The model is developed and maintained jointly by the Institute for Marine and Atmospheric Research Utrecht (IMAU, The Netherlands), the Joint Research Centre (JRC, Italy), the Royal Netherlands Meteorological Institute (KNMI), the Netherlands Institute for Space Research (SRON), and the NOAA Global Monitoring Laboratory (GML).

In CarbonTracker, TM5 separately simulates advection, deep and shallow convection, and vertical diffusion in both the planetary boundary layer and free troposphere. The carbon dioxide concentrations predicted by CarbonTracker do not feed back onto these predictions of winds.

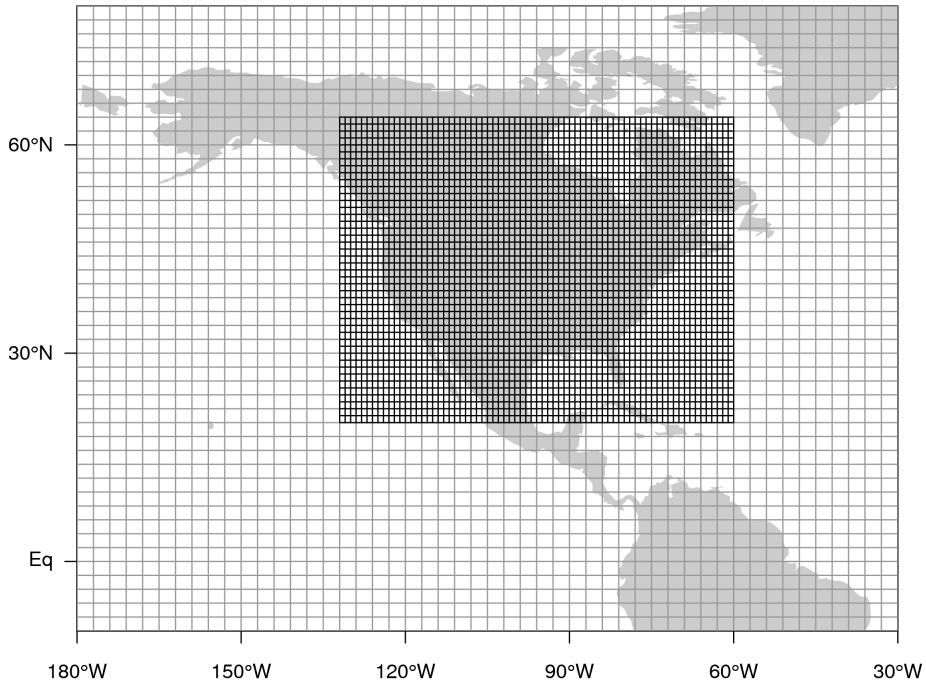


Figure 4.1: Nested grids used in CarbonTracker over North America. TM5 is a global model, but it employs nested grids to provide higher resolution over regions of interest. This figure shows the 1° × 1° nested regional grid over North America and a portion of the global 3° longitude × 2° latitude grid.

Prior to use in TM5, ECMWF meteorological data are preprocessed into coarser grids, with attention to retrieving a flow that conserves tracer mass. Like most numerical weather prediction models, advection in the parent ECMWF model is not strictly mass-conserving, so this step is crucial. In CarbonTracker, TM5 is currently run at a global 3° longitude × 2° latitude resolution with a nested regional grid over North America at 1° × 1° resolution (Figure 4.1). TM5 uses a dynamically-variable time step with a maximum length of 60 minutes. This overall timestep is dynamically reduced to maintain numerical stability, generally during times of high wind speeds. The timestep is divided in half and individual advection, diffusion, convection, and chemistry operators are applied symmetrically in each half step. Furthermore, transport operators in nested grids are modeled at shorter timesteps, so processes at the finest scales are conducted at an effective timestep of one-quarter the overall timestep. See Krol et al. (2005) for details.

The winds which drive TM5 come from the ERA5 reanalysis implemented in the European Centre for Medium-Range Weather Forecasts (ECMWF) modeling system (Hersbach et al., 2020). The ERA5 reanalysis uses CY41R2 version of the ECMWF Integrated Forecast System (IFS) model. That model uses a 12-minute time step and a spectral T639 horizontal resolution, which corresponds to approximately 28 km spacing at the equator on a reduced Gaussian grid. This version of the IFS has 137 model layers in the vertical, of which TM5 uses a 34-layer subset. The mean heights for these 34 levels are listed in Table 4.1.

4.2 Improvement in convective transport

Before 2013, TM5 was known to have difficulties representing the global surface distribution of sulfur hexafluoride (SF_6, see Figure 4.2 and Peters et al. (2004)). SF_6 is a nearly inert tracer in the atmosphere, with

Model Level	Mean Height (m)	Model Level	Mean Height (m)
1	33	18	9400
2	109	19	10131
3	255	20	11011
4	477	21	11749
5	814	22	12492
6	1273	23	13393
7	1835	24	14304
8	2556	25	15226
9	3315	26	16322
10	4205	27	17446
11	5026	28	18459
12	5603	29	20380
13	6186	30	24376
14	6771	31	29834
15	7355	32	35623
16	8086	33	42602
17	8816	34	123210

Table 4.1: Mean mid-level heights in meters above ground from the ERA5 reanalysis using the TM5 34 level subset.

very small surface and atmospheric sinks and an atmospheric lifetime of more than about 800 years. As a result of this long lifetime, its global budget is very well known from observations alone. SF_6 is thought to be released mainly via leakage from electrical transformers. Since the electrical distribution system is closely tied to fossil fuel consumption, SF_6 is often considered an analog for fossil fuel CO_2 in the atmosphere. It is useful for understanding the rate at which Northern Hemisphere land surfaces are ventilated to the free troposphere, and the rate of interhemispheric exchange in models (Patra et al., 2011).

As a result of more than a decade's worth of work on understanding the apparently sluggish mixing in TM5 as revealed by SF_6 simulations, a fault in one of the vertical mixing parameterizations of the model was discovered. When it was originally created, TM5 implemented the same planetary boundary layer (PBL) mixing and convection schemes as the parent ECMWF model. Recent comparisons between TM5, the ECMWF parent model, and radiosonde profile data show that the PBL scheme in TM5 performs similarly to that of the parent ECMWF model. The pre-2013 convective scheme, however, did not produce similar results in TM5 as compared to the ECMWF model.

Subsequent to the discovery of this convective transport issue, TM5 was modified to use parent model ERA5 convective fluxes directly. Using the parent model convective fluxes result in a significantly better SF_6 simulations. Simulations with these parent-model convective fluxes are said to use the "convective flux fix". Simulations with the convective flux fix show significantly improved agreement with SF_6 observations (see Figure 4.2).

When the convective flux fix was instituted in CT2013B, it resulted in the largest realignment of surface CO_2 fluxes in the history of the CarbonTracker program (Schuh et al., 2019). This is a prominent example of the sensitive reliance of atmospheric inversions on accurate atmospheric transport.

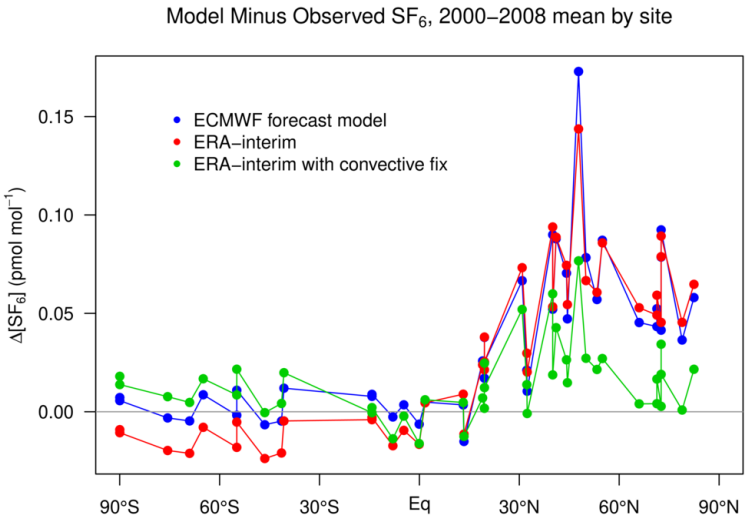


Figure 4.2: Long-term mean model residuals of SF_6 concentrations as a function of latitude. Residuals are defined as model-minus-observation, so a positive residual indicates the model has too much SF_6. Three different transport model simulations are shown. The ECMWF forecast (blue) and ERA-interim (red) transport simulations do not include the 2013 "convective flux fix". The ERA-interim with this convective flux fix is shown in green. Units are pmol mol^{-1}, or parts per trillion. CT2026 uses ERA5 transport with the convective flux fix.

Chapter 5

Observations

The observations of atmospheric CO_2 mole fraction made by NOAA GML and partner laboratories are at the heart of CarbonTracker. These measurements inform us on changes in the carbon cycle, whether those changes have regular cycles (such as the annual cycle of growth and decay of leaves and other plant matter), or irregular (such as the release of tons of carbon by wildfire). The results in CarbonTracker depend directly on the quality, location, and frequency of available observations. The level of detail at which we can retrieve information on the carbon cycle increases strongly with the density of the CO_2 observing network.

5.1 The CarbonTracker observational network

Observations simulated by CT2026 are supplied by the GLOBALVIEW+ data product version 11.0 (Schuldt et al., 2026), available at the NOAA GML ObsPack web site. This study uses measurements of air samples collected at 679 sites around the world by 80 laboratories:

- The Pennsylvania State University, Department of Meteorology and Atmospheric Science (PSU)
- NOAA Global Monitoring Laboratory (NOAA)
- Instituto de Pesquisas Energeticas e Nucleares (IPEN)
- Environment and Climate Change Canada (ECCC)
- AVOCET Group @ NASA LaRC (NASA-LaRC)
- National Institute for Environmental Studies (NIES)
- Commonwealth Scientific and Industrial Research Organisation, Environment (CSIRO)
- NASA Ames Research Center (NASA-Ames)
- Chinese Academy of Meteorological Sciences (CMA)
- National Institute for Space Research (INPE)
- Scripps Institution of Oceanography (SIO)
- Scripps Institution of Oceanography CO2 Program (SIO_CO2)
- Max Planck Institute for Biogeochemistry (MPI-BGC)
- Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvett (LSCE)
- Japan Meteorological Agency (JMA)
- NOAA Chemical Sciences Division (NOAA-CSD)
- National Institute of Water and Atmospheric Research (NIWA)
- Integrated Carbon Portal - RI (ICOS RI)

- ICOS Atmosphere Thematic Centre (ICOS-ATC)
- Norwegian Institute for Air Research (NILU)
- University of Bern, Physics Institute, Climate and Environmental Physics (KUP)
- Atmospheric Chemistry Research Group School of Chemistry University of Bristol (UNIVBRIS)
- Department of Earth and Planetary Sciences, Harvard University (HU)
- Earth Networks, Inc., an AEM company (EN)
- Netherlands Organisation for Applied Scientific Research (TNO)
- California Institute of Technology, Division of Geological and Planetary Science (CALTECH)
- Institute of Atmospheric Sciences and Climate (CNR-ISAC) (CNR-ISAC)
- Meteorological Research Institute (MRI)
- South African Weather Service (SAWS)
- Laboratoire d'Aérologie (LA)
- National University of Ireland - Galway (NUI-Galway)
- University of Exeter, Centre for Environmental Data Analysis (CEDA)
- Institut de Ciencia i Tecnologia Ambientals, Universitat Autonoma de Barcelona (ICTA-UAB)
- Lawrence Berkeley National Laboratory and ARM Climate Research Facility (LBNL-ARM)
- University of Maryland, College Park (UMD)
- NOAA Air Resources Laboratory (NOAA-ARL)
- Deutscher Wetterdienst, Hohenpeißenberg Meteorological Observatory (DWD)
- NASA Goddard Space Flight Center (NASA-GSFC)
- Korea Meteorological Administration, Korea Global Atmosphere Watch Center (KMA)
- Joint Research Unit Ecology of Guianan Forests (EcoFoG)
- University of East Anglia (UEA)
- National Center For Atmospheric Research (NCAR)
- University of Heidelberg, Institut fuer Umweltphysik (UHEI-IUP)
- National Physical Laboratory (NPL)
- Hong Kong Observatory (HKO)
- Lund University - Centre for Environmental and Climate Research (LUND-CEC)
- Hungarian Meteorological Service (HMS)
- Karlsruhe Institute of Technology (IMK-ASF) (KIT/IMK-ASF)
- Institute for Atmospheric and Environmental Sciences, University of Frankfurt (IAU)
- Joint Research Centre (JRC)
- Izana Atmospheric Research Center, Meteorological State Agency of Spain (AEMET)
- Estonian University of Life Sciences (EMU)
- Swiss Federal Laboratories for Materials Science and Technology (EMPA)
- AGH University of Krakow (AGH)
- University of Minnesota (UofMN)
- Global Change Research Institute of the Czech Academy of Sciences (CAS)
- Integrated Carbon Observation System - Flask and Calibration Laboratory (ICOS)
- University of Wisconsin (UofWI)
- National Agency for New Technology, Energy, and Environment (ENEA)
- University of Groningen (RUG), Centre for Isotope Research (CIO) (RUG)
- Groupe de Spectrométrie Moléculaire et Atmosphérique (GSMA)
- Kenya Meteorological Department (KMD)
- Royal Holloway University London (RHUL)
- Oregon State University (OSU)
- Laboratory for Atmospheres, Environments, Space Observations (LATMOS) (LATMOS)
- Finnish Meteorological Institute (FMI)

Source	Online availability	Period of use
GLOBALVIEW+ v11.0	Available via ObsPack download	2000-2024
NRT 11.0	Available via ObsPack download	2025
Manaus profiles	Available via ObsPack download	2017-2024

Table 5.1: Sources of CO_2 observational data for CT2026

- Ricerca sul Sistema Energetico (RSE)
- Royal Belgian Institute for Space Aeronomy (BIRA-IASB)
- Savannah River National Laboratory (SRNL)
- University of Helsinki (UHELS)
- Aarhus University (AU)
- University of Virginia (UofVA)
- Umweltbundesamt Offenbach/University of Heidelberg, Institut fuer Umweltphysik (UBA/UHEI-IUP)
- Lawrence Berkeley National Laboratory (LBNL)
- Forest Ecology and Management, SLU Umeå (SLU)
- Center for Atmospheric and Oceanic Studies, Tohoku University (TU)
- National Institute of Polar Research (NIPR)
- Main Geophysical Observatory (MGO)
- Utah Atmospheric Trace gas & Air Quality (U-ATAQ)
- Umweltbundesamt Offenbach (UBA-Germany)

The CO_2 measurement data assimilated in CT2026 are freely available for download from the GML ObsPack web portal or from partner websites. The bulk of assimilated measurements come from the GLOBALVIEW-plus v11.0 product. Additional observations were gathered from specialized ObsPack products as detailed in Table 5.1.

We also make available an ObsPack containing the simulated values of all measurement data considered by CT2026. This CT2026 ObsPack contains most, but not all, of the measured values. Measured values are only distributed directly when we have permission to do so.

Users are encouraged to review the usage requirements for these data products, and to contact the measurement laboratories directly for details about the observations.

With the advent of GLOBALVIEW+ in 2015, data are now presented to CarbonTracker with a higher temporal frequency than in past observational products. At sites with quasi-continuous monitoring, CT2026 assimilates hourly average CO_2 concentrations. In the past, a single daily assimilation value was constructed at these sites, generally a four-hour average during well-mixed background conditions. At continental sites, this four-hour period was generally from local noon to 4pm; at many mountain sites background conditions are met at nighttime when upslope winds are uncommon. Using GLOBALVIEW+, CarbonTracker can now assimilate each hourly average during these background conditions independently. For many sites, all available hourly averages throughout the day are assimilated. Details vary by dataset, but can be checked at the interactive data plotting page.

Note that all of these observations are calibrated against the same world CO_2 standard (WMO-X2019).

Starting with GLOBALVIEW+, we generally use the recommendations of data providers as to which observations are appropriate for assimilation. Such observations are identified by a variable in the ObsPack distribution, obs_flag. Only observations with obs_flag = 1 are identified for assimilation by data providers. We modify the designation of assimilation data for Environment and Climate Change Canada

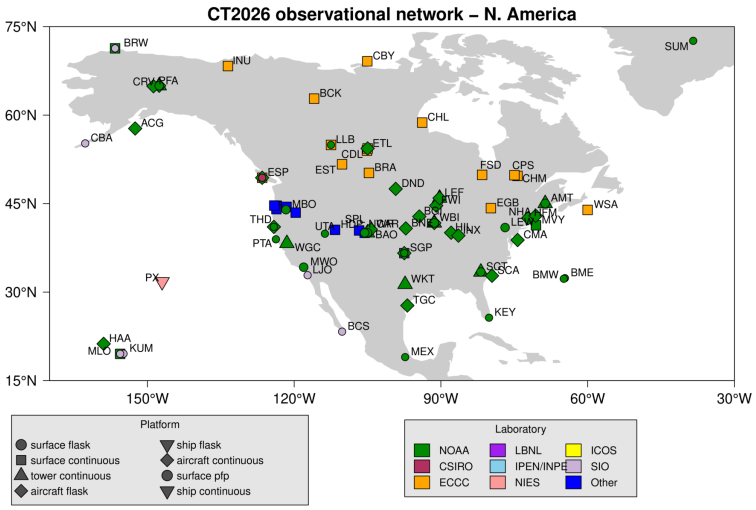


Figure 5.1: CarbonTracker observational network over North America. See the CarbonTracker interactive network map for more details.

quasi-continuous sampling sites. For these data, obs_flag is set to 1 by the data provider for times when they represent the daily minimum CO_2 concentration. This is generally later in the day than our standard scheme of local noon-4pm used to represent times of well-mixed PBLs. For these datasets, we have changed obs_flag to indicate assimilation only for the local noon-4pm time period. These selected observations are further filtered based on the CCG curve fitting routine of Thoning et al. (1989). This filter fits a smooth curve to the selected observations, and measurements more than 3 standard deviations away from this curve are excluded from assimilation.

At mountain-top sites (*e.g.* MLO, NWR, and SPL), it is usually nighttime hours that are selected for assimilation, as these tend to be the most stable time period. Nighttime hours also avoid periods of upslope flows that contain local vegetative and/or anthropogenic influence.

We assimilate CO_2 measurements from NOAA light aircraft profiling time series, from intakes at multiple levels on NOAA tall towers, and from extensive shipboard and Siberian tower measurements collected by our partners at NIES. These datasets can be explored at the interactive data plotting page.

We apply a further selection criterion during the assimilation to exclude non-marine boundary layer (MBL) observations that are very poorly forecasted in our framework. We use the so-called model-data mismatch in this process, which is the random error ascribed to each observation to account for measurement errors as well as modeling errors of that observation. We interpret an observed-minus-forecasted mole fraction that exceeds 3 times the prescribed model-data mismatch as an indicator that our modeling framework fails. This can happen for instance when an air sample is representative of local exchange not captured well by our $1^\circ \times 1^\circ$ fluxes, when local meteorological conditions are not captured by our offline transport fields, but also when large-scale CO_2 exchange is suddenly changed (*e.g.* fires, pests, droughts) to an extent that can not be accommodated by our flux modules. This last situation would imply an important change in the carbon cycle and has to be recognized by the researchers when analyzing the results. In accordance with the 3-sigma rejection criterion, about 0.2% of the observations are discarded through this mechanism in our assimilation.

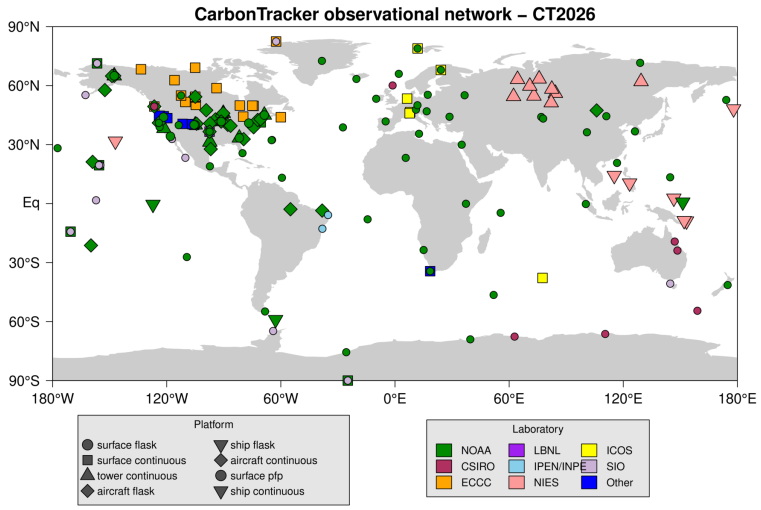


Figure 5.2: CarbonTracker global observational network. See the CarbonTracker interactive network map for more details.

5.2 Adaptive model-data mismatch

The statistical optimization method we use to constrain surface CO_2 fluxes requires that each assimilation constraint is assigned a "model-data mismatch" (MDM) error value. This is meant to express the statistics of simulated-minus-observed CO_2 observations we could expect if CarbonTracker were using perfect surface fluxes. Such deviations arise from many sources, including random noise in the measurement system, *in situ* variability that we do not expect to resolve in our model, and faults with the atmospheric transport model. Generally, transport and inverse model faults are the dominant terms in MDM values. The MDM is one of two major "tuning knobs" used to adjust the performance of our ensemble Kalman filter. The other is also an error quantity, meant to represent the expected error on our first-guess fluxes. Discussion of this prior covariance error can be found in Section 6.3.1.

Prior to CT2015, CarbonTracker used a single MDM value for each assimilation dataset. The NOAA continuous observations at the 396m level of the WLEF tower in northern Wisconsin, for example, were assigned a MDM of 3.0 ppm, meaning that the residuals between model-forecasted measurements and the actual observed concentrations are expected to be unbiased (*i.e.*, have a mean of zero) and have a standard deviation of 3 ppm. In practice, however, we have found that it is far easier to simulate wintertime observations than those during summer. This is mainly due to higher ambient variability of CO_2 in the summer.

Starting with CT2016, we began to use an empirical scheme to assign MDM values, exploiting statistics of model performance from independently-configured preliminary inversions. The posterior residuals for each dataset are classified into relevant bins, and then statistics of model performance are analyzed within each of those bins. For every dataset, these bins include equally-spaced intervals of one-tenth of a year. For analyzers collecting data throughout the day, we also classify the measurements into 4-hour intervals of local time. For aircraft datasets, we further classify measurements into vertical levels of 1000m thickness (0-1000 m ASL, 1000-2000 m ASL, etc.). For each of these bins, bias and random error are combined to form total deviation from observed values as a root-mean square error (RMSE). The assigned MDM is set to a constant fraction of this total RMSE. This scaling is meant to force the assimilation scheme to extract as much information as possible from available observations. We use two different scaling factors to convert RMSE to MDM, depending on whether the preliminary inversions actually assimilated the measurements in

the relevant bin, or merely simulated those measurements. For measurements assimilated by the preliminary inversions, the MDM is 0.95 $*$ RMSE. For measurements not assimilated in the preliminary inversions, the MDM is 0.85 $*$ RMSE.

The adaptive MDM scheme performs well in terms of average χ^2, which in an optimally-tuned system should be close to 1.0 for each dataset (see Table B.1). Notably, the seasonal variations of MDM successfully compensate for the higher ambient variability of CO_2 at continental sites during the growing season. It is, however, an iterative process, requiring that we conduct a previous inversion. For various reasons, this previous inversion performed before CT2026 differs in significant aspects from the actual CT2026 inversions. These differences have led to MDM values which are slightly too large and thus average χ^2 values which are generally smaller than the target of 1.0 (in some cases, as low as 0.2 or 0.3). The next iteration of CarbonTracker will be able to use the more recent CT2026 inversions to refine the adaptive MDM scheme.

Duplicate observations are identified as those within 50 minutes temporally, 10m vertically, and 0.05 degrees of latitude and longitude laterally (nominally, about 5km). The MDM for such observations is inflated by $\sqrt{n}$, where n is the number of duplicates.

Chapter 6

Ensemble data assimilation

Data assimilation is the process by which a model simulation is adjusted to agree with observations. Model simulations may drift off from reality for a number of reasons. Some models are highly nonlinear, and depend sensitively on knowing the system state with high accuracy. Weather models fall into this category, and as a result reliable forecast systems depend on having a constant stream of meteorological data to correct their simulations. In contrast, models like CarbonTracker need data assimilation not because the controlling dynamics are nonlinear, but because those dynamics are not well known. CarbonTracker uses approximate or estimated rules about the evolution of surface CO_2 fluxes, then corrects these approximate projections using observational constraints. The resulting optimal surface flux estimates can then be used to better understand the functioning of the carbon cycle.

Data assimilation is usually a cyclical process, in which estimates get refined over time as more observations become available. Mathematically, data assimilation can be performed using a wide variety of techniques, including variational and ensemble methods. Assimilation systems involving simulations of the global atmosphere are often implemented on highly parallel supercomputers in order to distribute the workload among many computational cores. CarbonTracker is an example of such a model because it relies heavily on estimates of global atmospheric transport.

CarbonTracker model predictions are limited by the relatively simple representations of CO_2 surface exchange used to predict land biosphere and ocean fluxes and emissions from fossil fuel combustion and wildfires. As described in the following section, we use data assimilation techniques to modify these surface fluxes so that the resulting atmospheric distribution of CO_2 agrees optimally with measurements. We do this by estimating a set of spatially- and temporally-varying scaling factors that multiply first-guess predictions from prior flux models. Data assimilation allows us to determine optimal values for these scaling factors.

6.1 Parameterization of unknowns

CO_2 fluxes $F(x, y, t)$ in CarbonTracker are parameterized according to

$$F(x, y, t) = (1 + \lambda(x, y, t))\Big(F_{\text{land}}(x, y, t) + F_{\text{ocean}}(x, y, t)\Big) + F_{\text{FF}}(x, y, t) + F_{\text{fire}}(x, y, t), \tag{6.1}$$

where F_{land}, F_{ocean}, F_{FF}, and F_{bio} are prior flux model predictions for land biosphere, ocean, fossil fuel and wildfire emissions respectively, and $1 + \lambda$ represents a set of unknown multiplicative scaling factors

applied to the fluxes, to be estimated in the assimilation. These scaling factors are the final product of our assimilation and together with the prior flux models determine CarbonTracker optimized fluxes. Note that no scaling factors are applied to the fossil fuel and fire modules. The fossil fuel and wildfire fluxes are relatively well-known from prior flux models compared to highly-uncertain land biosphere and ocean fluxes, and as a result we impose those emissions without modification in our model.

6.1.1 Optimization regions

The scaling factors $1 + \lambda$ are estimated independently for each week and optimization region. They are assumed to be constant over this time period and spatial domain. Each scaling factor is associated with a particular region of the globe, as in the Transcom inversion study (Gurney et al., 2002). Currently the geographic distribution of these optimization regions is fixed. The choice of regions is a strong *a priori* design decision determining the reliability of the resulting fluxes. In particular, the scale of optimization regions is chosen to minimize "aggregation errors" (Kaminski et al., 2001), while limiting the set of unknown parameters to a manageable number. Following Jacobson et al. (2007), we have divide the global ocean into 30 basins encompassing large-scale ocean circulation and biogeochemical features. The terrestrial biosphere is divided up according to ecosystem type and geographical domain. Specifically, each of the 11 Transcom land regions is subdivided into a maximum of 19 "ecoregions" according to its Olson et al. (1992) vegetation classification. The set of ecoregions over North America is summarized in Table 6.1 and Figure C.3. Note that there is currently no requirement for ecoregions to be contiguous, and a single scaling factor can be applied to the same vegetation type on both sides of a continent. Further details on ecoregions can be found in Section C.

Theoretically, this approach leads to a total number of 11*19+30=239 optimizable scaling factors for each week, but the actual number of optimization regions is only 156 since some ecosystem types are not represented in every Transcom region. It should be noted also that we have chosen to not optimize scaling factors for ice-covered regions, inland water bodies, and deserts, since the CO_2 flux from these regions is negligible.

It is important to note that even though only one parameter is available to scale, for instance, the flux from coniferous forests in Boreal North America, each $1^\circ \times 1^\circ$ grid box predominantly covered by coniferous forests will have a different optimized flux $(1 + \lambda)F_{\text{land}}(x, y, t)$ depending on local temperature, radiation, and emissions as simulated by the prior flux model.

Ecosystem types are based on the vegetation classification of Olson et al. (1992). Note that we have adjusted the original 29 categories into only 19 regions. This was done mainly to fill the unused categories 16, 17, and 18, and to group the similar categories 23-26 and 29. Table 6.1 shows each vegetation category considered. Percentages indicate the relative area in North America associated with each category.

Each $1^\circ \times 1^\circ$ pixel of our domain was assigned one of the categories above based on the Olson category that was most prevalent in the $0.5^\circ \times 0.5^\circ$ underlying area.

6.1.2 Assimilation window

Measured CO_2 mole fractions are the result of upstream surface fluxes and atmospheric transport, which includes both advective movement and diffusive mixing. Near-field surface fluxes can cause significant changes in CO_2 mole fractions, whereas flux signals from further upstream become spread out and diluted. Generally speaking, the longer in the past a flux event occurred, the smaller its impact will be on a given

Category	Olson V 1.3	Percentage area
1	Conifer Forest	19.0%
2	Broadleaf Forest	1.3%
3	Mixed Forest	7.5%
4	Grass/Shrub	12.6%
5	Tropical Forest	0.3%
6	Scrub/Woods	2.1%
7	Semitundra	19.4%
8	Fields/Woods/Savanna	4.9%
9	Northern Taiga	8.1%
10	Forest/Field	6.3%
11	Wetland	1.7%
12	Deserts	0.1%
13	Shrub/Tree/Suc	0.1%
14	Crops	9.7%
15	Conifer Snowy/Coastal	0.4%
16	Wooded tundra	1.7%
17	Mangrove	0.0%
18	Non-optimized areas (ice, polar desert, inland seas)	0.0%
19	Water	4.9%

Table 6.1: Ecosystem types over North America

sample of air (although it will be spread out through a larger volume of the atmosphere). Thus we choose an "assimilation window" that represents how far back in time we expect to be able to pinpoint a given flux signal from available measurements. A good discussion of this topic can be found in Bruhwiler et al. (2005).

In previous versions of CarbonTracker, the assimilation window was chosen to be five weeks long, meaning that a measurement could cause revisions in surface fluxes only over the 5 weeks leading up to that measurement. In CT2026, we have extended the assimilation window length to 12 weeks. This helps to resolve fluxes in regions of the world with less dense observational coverage (the tropics, Southern Hemisphere, and parts of Asia).

This assimilation window is moved forward on each cycle of our estimation system, so that new weeks are introduced at the "head" of the filter, and the weeks that fall out the "tail" of the filter are finalized. Prior to CT2026, the 5-week assimilation window was moved forward one week at a time. In CT2026, the 12-week assimilation window is moved forward two weeks at a time. Scaling factors $1+\lambda$ retain their weekly resolution. Each cycle of the inversion system requires running the atmospheric model for a length of time equal to the assimilation window length plus the window step size. For previous CarbonTracker releases, this was 6 weeks per cycle; for CT2026 it is 14 weeks per cycle. The extra computing time required by the longer assimilation window is balanced somewhat by the two-week stepping.

Ensemble size and localization

The ensemble system used to solve for the scalar multiplication factors is similar to that in Peters et al. (2005) and based on the square root ensemble Kalman filter of Whitaker and Hamill (2002). Ensemble

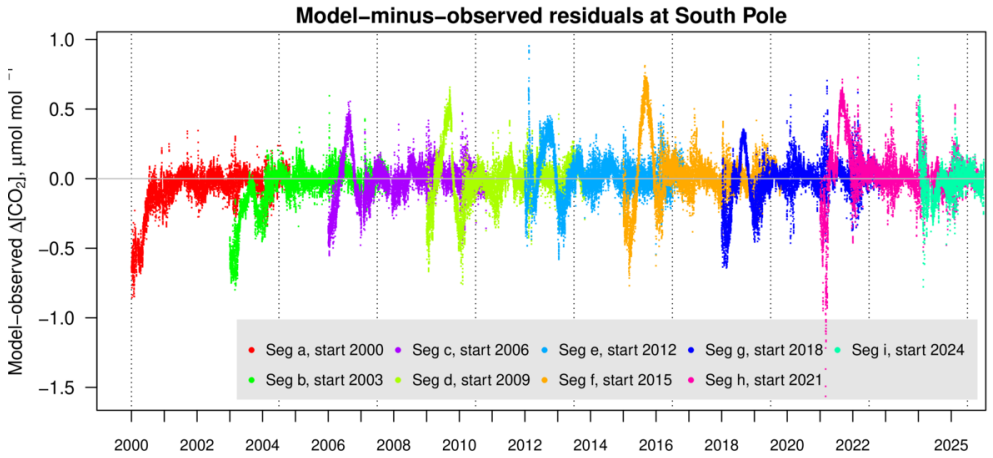


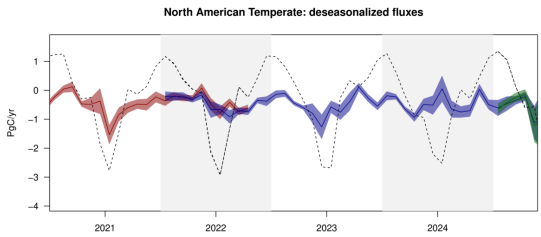
Figure 6.1: Time series of model residuals from observations at the South Pole from the individual segments of the p7 simulation. Each of the nine segments, coded as "Seg a" through "Seg i", is shown in a different color (see legend). Each of the times at which CT2026 transitions from one segment to the next is shown with a vertical dotted line. Each segment generally starts with large residuals against measurements, but after 18 months of spinup, overlapping segments have consistent residual statistics. Being as remote as it is, the South Pole is very sensitive to segment overlap problems.

statistics are created from 600 randomly-chosen members, each with its own background CO_2 concentration field to represent the time history of that member's surface fluxes. The ensemble Kalman filter looks for correlations between these random flux perturbations and resulting changes in simulated CO_2 measurements. We might expect that the entire ensemble would agree that increasing the CO_2 flux in a given region results in greater simulated CO_2 at a nearby downwind site. However, because we approximate the flux covariance matrix with a random sample of 600 members, sometimes spurious correlations appear. It is unphysical, for instance, that a measurement at Summit, Greenland could be strongly influenced by surface exchange in the southern Indian Ocean, within the time span of our assimilation window. Any such correlation between the flux ensemble and the measurement in question is likely spurious. Localization is a technique developed for numerical weather prediction in which unphysical correlations are diagnosed and systematically ignored (Houtekamer and Mitchell, 1998). We perform localization for our measurements, but only for certain datasets. Notably, it is not used for datasets judged to represent hemisphere-scale signals, such as those from marine boundary layer sites in remote locations.

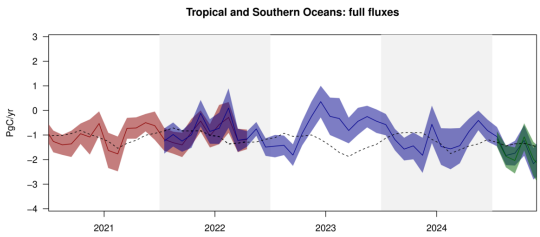
Our localization technique is based on the linear correlation coefficient between the 600 parameter deviations and 600 observation deviations for each parameter/observation pair. If the relationship between a parameter deviation and its modeled observational impact is statistically significant, then that relationship is retained. Otherwise, the relationship is assumed to be spurious noise due to the numerical approximation of the covariance matrix by the limited ensemble. We accept relationships that reach 95% significance in a Student's T-test with a two-tailed probability distribution.

6.2 Physical parallelization

In CT2026 we break our analysis period into segments initialized every three years, which are then run simultaneously. These segments start on 1 January of 2000, 2003, . . . , 2021, and 2024. Each segment is initialized with an analyzed CO_2 field appropriate to the starting date and the same prior state (*cf.* Section 6.3).



(a) Segment flux overlap for the North America temperate Transcom land region.



(b) Segment flux overlap for tropical and southern hemisphere oceans.

Figure 6.2: Segment overlap in flux space for 2021-2025. Fluxes for each segment are shown with colored lines and one-sigma uncertainties shaded around those lines. Unoptimized fluxes are shown with a dashed line. All fluxes are in PgC/yr for the specified region. Spin-up months at the beginning of each segment are omitted. (a) Overlap for the North America temperate Transcom land region. A reference seasonal cycle (the CT2022 climatology for this region) has been removed from all flux components. (b) Overlap for ocean regions in the tropics and southern hemisphere. Unlike Figure 6.2a, no referrence seasonal cycle was removed.

"Physical parallelization" of this type was introduced for CO_2 analyses by Chevallier (2013). This procedure has reduced the time to produce a CarbonTracker release by about an order of magnitude.

The final CT2026 fluxes and simulated mole fractions are formed by stitching together the parallelized segments. A given segment n takes up to a 18 months to equilibrate in mole fraction and flux space to the same state as the previous segment $n-1$, which has integrated for three additional years. Chevallier (2013) evaluated segment overlap in a variational framework and found that a three-month overlap was appropriate for that system. With the ensemble Kalman filter, we can evaluate the segment overlap not only in simulated CO_2 mole fraction space (Figure 6.1), but also in flux and flux uncertainty space (Figure 6.2). We find in simulation experiments and in the current operational inversions that an overlap of 18 months produces statistically consistent segment states. Thus the CT2026 results are composed of first-segment data from January 2000 to mid-July 2004, at which point results from the second segment (initialized at January 1, 2003) are imposed. This process is continued to stitch together results from all nine segments.

6.3 Dynamical model

With CT2026 we introduce a new dynamical model, based on a linear time propagation operator, Ψ. Previous versions of CarbonTracker used a more *ad hoc* dynamical model that was not consistent with the mathematics of the Kalman filter. The time propagation operator is used to estimate the prior state at time t, denoted $\{\lambda^-[t], P_\lambda^-[t]\}$, from the posterior state at time $t-1$, which is written as $\{\lambda^+[t-1], P_\lambda^+[t-1]\}$. This is a linear transformation of the state, expressed as

$$\lambda^-[t] = \Psi\lambda^+[t-1] + \epsilon_\Psi, \tag{6.2}$$

and

$$P_\lambda^-[t] = \Psi^{\mathrm{T}} P_\lambda^+[t-1]\Psi + P_\Psi, \tag{6.3}$$

where the prior value of the scaling factors for timestep t is $1+\lambda^-[t]$, the posterior at timestep $t-1$ is $1+\lambda^+[t-1]$, and ϵ_Ψ is a mean-zero noise process representing uncertainty in this temporal propagation model. The covariance matrix of the ϵ_Ψ noise process is P_Ψ.

In practice the mean state is evolved without adding an explicit ϵ_Ψ, but the ensemble deviations are seeded with random noise drawn from P_Ψ. This process noise is extremely important in the Kalman filter, since it is responsible for maintaining a minimum amount of error covariance in the ensemble. In effect this keeps freedom, or "wiggle-room", in the state so that the system can respond to new signals in measurements destined for assimilation. This state freedom prevents premature convergence to a faulty solution. The addition of process noise is a sort of error inflation.

The time propagation scheme chosen for CT2026 is based on modified persistence. In this scheme, scaling factor parameters are assumed to decay back towards unity with a certain smoothing timescale. Mathematically, we express this as:

$$\Psi[i,i] = \exp \frac{-\Delta t}{\tau}, \tag{6.4}$$

where τ is a timescale parameter and Δt is the amount of time over which the state is being propagated. We use $\tau = 7$ days. Ψ is a diagonal matrix, which means that the scaling factors evolve independently from one another.

6.3.1 Structure of master prior covariance

The "master" prior covariance matrix P_0^- is used to initialize the ensemble Kalman filter on its first timestep, and to inject process noise in the time propagation step. It describes the magnitude of the uncertainty on each parameter, plus their correlations. The correlations between the same ecosystem types in different Transcom regions decrease exponentially with distance scale $L = 2000$km, and thus assumes a coupling *between* the behavior of the same ecosystems in close proximity to one another (such as coniferous forests in Boreal and Temperate North America). Furthermore, all ecosystems *within* tropical Transcom regions are coupled decreasing exponentially with distance since we do not believe the current observing network can constrain tropical fluxes on sub-continental scales, and want to prevent spurious compensating source/sink pairs ("dipoles") to occur in the tropics.

While the correlation structure discussed above has remained fixed in all CarbonTracker releases, we have been changing the overall magnitude of the covariance matrix in an attempt to mitigate seasonal biases in our simulated CO_2 fields. Since those biases appear mostly in comparison to measurement data over land, and also because the annual cycle of CO_2 in the atmosphere is dominated by the terrestrial carbon cycle, we experimented by loosening the land prior constraint. As it turns out, these biases were mostly due to our original short assimilation window length (Section 6.1.2) and the convective flux problem discussed in Section 4.2. As a result, for CT2026 we were able to scale back the land prior covariance. "L-curve" analysis (Hansen, 1998) suggested that these covariances could be reduced even from the original levels, and new values are shown in red in the lower panel of Figure 6.3.

6.3.2 Process Noise

As previously discussed, the process noise P_Ψ is a crucial "tuning knob" in the Kalman filter. While formally it represents uncertainty in the time propagation methodology, in practice it is responsible for maintaining a minimum level of flexibility in the state. In a series of simulation experiments, different versions of the CarbonTracker dynamical model Ψ and P_Ψ were tested until an optimal version was identified.

In these simulation experiments, synthetic measurement data were generated from a known truth condition. These synthetic observations were then corrupted with random noise drawn from the assumed model-data mismatch (see Section 5.2). Each simulation experiment consisted of a three-year inversion in which these

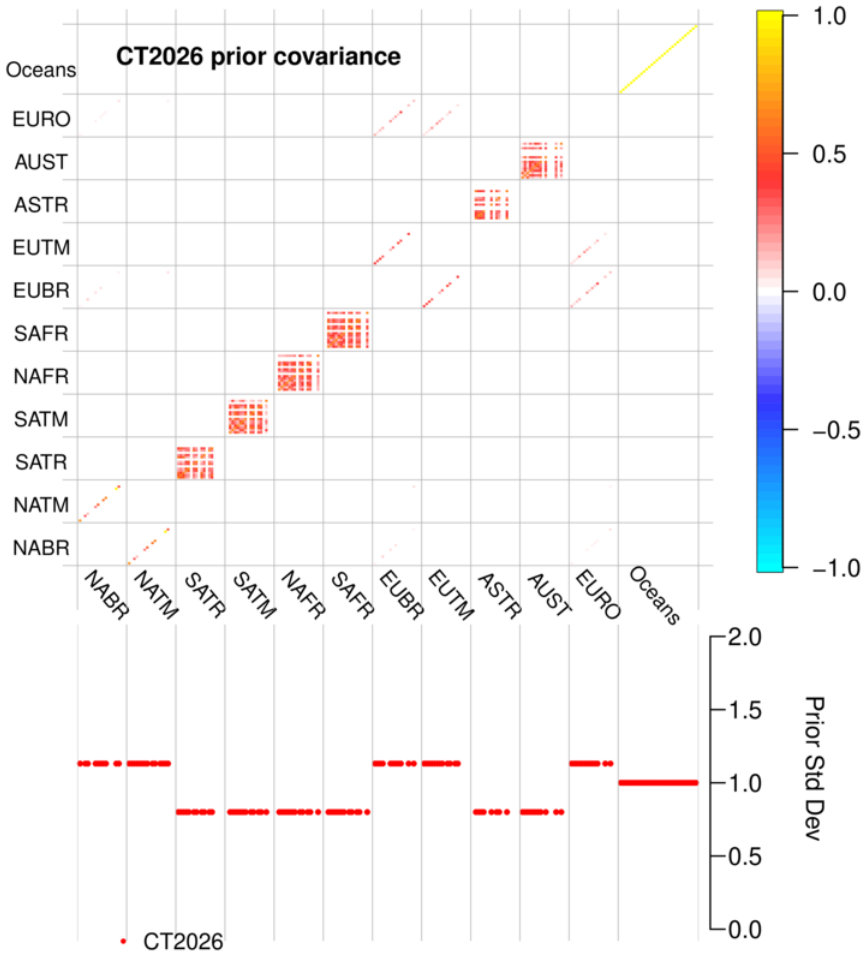


Figure 6.3: CT2026 prior covariance structure. The prior covariance matrix (top panel) and the square root of diagonal members of this matrix (bottom panel). Covariance matrix quantities are dimensionless squared scaling factors, and the bottom panel is the square root of this. Transcom land regions form the first 11 large divisions on the axes here. As described in Section C, each of those regions contains 19 potential ecosystems. Correlations between similar ecosystems in proximate Transcom regions are visible in North America (*e.g.* NABR and NATM, the boreal and temperate North American regions) and Eurasia. Within tropical Transcom regions, however, differing ecosystems are assigned a non-zero prior covariance, which is visible here as red block-like structures on the diagonal within, for example, the South America Tropical (SATR) Transcom region. Ocean regions have a more complicated covariance structure that depends on which prior is used; the structure shown here is that of the ocean inversion flux prior. The lower panel of this diagram compares the on-diagonal elements of the prior covariance matrix by plotting their square roots. The resulting standard deviations are directly comparable to the percentages discussed in Section 6.3.1 above; 0.5 is equivalent to 50%. The retuning of the covariance matrix for CT2026 (in red) is made evident by also showing these values from previous CarbonTracker releases in light blue and black.

synthetic measurements were assimilated into a CarbonTracker run. These fluxes estimated by these runs were then evaluated against the truth condition. This evaluation consisted both of quantifying the differences between retrieved and true fluxes, and also evaluating whether those differences were spanned by the posterior uncertainties.

The most successful process noise parameterization was found by adding a scaled version of the initial prior covariance matrix P_0^-:

$$P_\Psi = 0.2 P_0^- \tag{6.5}$$

6.3.3 Posterior uncertainties in CarbonTracker

In CT2026, our error estimates on optimized fluxes are considerably more realistic as a result of adopting a formally-correct linear time propagation scheme. Notably, they reflect the density of measurement constraints in various regions. Previous CarbonTracker releases were far more aggressive about covariance inflation, which resulted in unrealistically large posterior errors.

CT2026 uncertainties do not resolve errors which are correlated in time. This is due to the limited extent of our assimilation window, in which these temporal error covariances are only modestly resolved. The present assimilation window of 12 weeks, with a two-week stepping, means that only the very shortest temporal correlations would be resolved. However, these short (anti) correlations are quite evident in CT2026 posterior fluxes, and we are exploring methods for including temporal information in CarbonTracker uncertainties.

Chapter 7

Statistical performance of CT2026

7.1 Measurement data

Starting with CT2022, we started to reserve about 5% of available assimilation data for a cross-validation exercise. To the extent possible, withheld measurements were chosen to be independent from other observations. For surface flask observations, which are generally collected on a weekly time basis, all samples are considered independent from one another, so withheld data were selected randomly. Aircraft flasks collected during a profile are considered co-dependent, so entire profiles were randomly selected for withholding. Finally, for *in situ* analyzers with quasi-continuous sampling (towers, observatories), 24-hour periods were randomly chosen and that entire day's worth of data were withheld. Shipboard quasi-continuous datasets, which account for more than a quarter of assimilation data for CT2026 were inadvertently excluded from this cross-validation exercise. As a result, about 5% (166,097 measurements out of almost 3.5 million) were withheld.

Each residual from the withheld measurements was divided by its prescribed model-data mismatch (MDM) error, to form a set of normalized residuals. From this sequence we can compute the mean (or reduced) chi-squared statistic $\overline{\chi^2}$, which for a perfect set of independent, normally-distributed variates should approach unity, is 1.2. This is quite similar to the same statistic for assimilated data, which is 1.04 (see Table 7.1).

These slightly inflated $\overline{\chi^2}$ values are actually as intended. They indicate that the model-data mismatch values used in CarbonTracker are too small. This in turn suggests that the RMSE deflation described is Section 5.2 probably is too agressive, but it also means that the Kalman filtering scheme will attempt to fit measurement data at the expense of retrieving a less-smooth flux solution. This overfitting will be revised in upcoming CarbonTracker releases.

Observation type	No. Obs.	Bias (ppm)	Standard Deviation (ppm)	$\overline{\chi^2}$
Assimilated	3 454 091	-0.07	3.2	1.04
Withheld	166 097	-0.07	3.47	1.24
Not Assimilated	5 923 015	-2.57	8.26	–

Table 7.1: Statistical performance of CT2026 for assimilated, withheld, and not assimilable measurements. Bias is computed as simulated minus measured value. Unassimilated measurement data generally do not have MDM values, so it is not possible to compute $\overline{\chi^2}$ statistics for them.

Chapter 8

Resources and References

- Global Fire Emissions Database (GFED) web page
- NASA GFED_CMS CASA project
- Olson ecosystem types
- GFED_CMS web page at NACP
- DOE Energy Information Administration (EIA)
- BP Statistical Review of World Energy
- EDGAR Database
- DMSP satellite nightlight data
- Centre for Air Transport and the Environment (CATE), AERO2k aviation emissions inventory
- The Surface Ocean CO_2 Atlas
- NOAA Pacific Marine Environmental Laboratory
- Ocean Acidification
- GML Carbon Cycle Program
- WMO/GAW Report No. 168, 2006

Bibliography

Andres, R. J., Boden, T. A., and Higdon, D. (2014). A new evaluation of the uncertainty associated with CDIAC estimates of fossil fuel carbon dioxide emission. *Tellus B: Chemical and Physical Meteorology*, 66(1):23616.

Andres, R. J., Gregg, J. S., Losey, L., Marland, G., and Boden, T. A. (2011). Monthly, global emissions of carbon dioxide from fossil fuel consumption. *Tellus B: Chemical and Physical Meteorology*, 63(3):309–327.

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibánhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S. (2016). A multi-decade record of high-quality fCO_2 data in version 3 of the surface ocean CO_2 atlas (SOCAT). *Earth System Science Data*, 8(2):383–413.

Basu, S., Lehman, S. J., Miller, J. B., Andrews, A. E., Sweeney, C., Gurney, K. R., Xu, X., Southon, J., and Tans, P. P. (2020). Estimating us fossil fuel CO_2 emissions from measurements of ^{14}C in atmospheric CO_2. *Proceedings of the National Academy of Sciences*, 117(24):13300–13307.

Basu, S., Miller, J. B., and Lehman, S. (2016). Separation of biospheric and fossil fuel fluxes of CO_2 by atmospheric inversion of CO_2 and $^{14}CO_2$ measurements: Observation system simulations. *Atmos. Chem. Phys*, 16(9):5665–5683.

Basu, S. and Nassar, R. (2026). Fossil fuel CO_2 emissions for the OCO2 Model Intercomparison Project (MIP).

Blasing, T., Marland, G., and Broniak, C. (2004). Estimates of monthly CO_2 emissions and associated $^{13}C/^{12}C$ values from fossil-fuel consumption in the USA (1981-2003). *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.*

Boden, T. A., Marland, G., and Andres, R. J. (2017). Global, regional, and national fossil fuel CO_2 emissions. Technical report, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

British Petroleum (2019). *BP Statistical Review of World Energy*. Number 68. BP p.l.c., London.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P. (2005). An improved Kalman smoother for atmospheric inversions. *Atmospheric Chemistry and Physics*, 5:2691–2702.

Caldeira, K. and Wickett, M. E. (2003). Anthropogenic carbon and ocean *p*H. *Nature*, 425(6956):365–365.

Chevallier, F. (2013). On the parallelization of atmospheric inversions of CO_2 surface fluxes within a variational framework. *Geoscientific Model Development*, 6(3):783–790.

Collatz, G., Ball, J., Grivet, C., and Berry, J. A. (1991). Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a lam inar boundary layer. *Agricultural and Forest Meteorology*, 54(2):107–136.

Collatz, G., Ribas-Carbo, M., and Berry, J. (1992). Coupled photosynthesis-stomatal conductance model for leaves of C4 plants. *Australian Journal of Plant Physiology*, 19(5):519–538.

Crippa, M., Guizzardi, D., Pagani, F., Schaaf, E., Muntean, M., et al. (2025). EDGAR (Emissions Database for Global Atmospheric Research) 2025 GHG release.

Dickson, A., Sabine, C., and Christian, J. (2008). Guide to best practices for ocean CO_2 measurements. *PICES special publication 3*, page 191.

Energy Institute (2025). *Statistical Review of World Energy 2025*. Energy Institute, London.

Erb, M. and Marland, G. (2026). CDIAC at AppState: global, regional, and national fossil-fuel CO_2 emissions.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Aas, K., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Bellouin, N., Benoit-Cattin, A., Berghoff, C. F., Bernardello, R., Bopp, L., Brasika, I. B. M., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Collier, N. O., Colligan, T. H., Cronin, M., Djeutchouang, L. M., Dou, X., Enright, M. P., Enyo, K., Erb, M., Evans, W., Feely, R. A., Feng, L., Ford, D. J., Foster, A., Fransner, F., Gasser, T., Gehlen, M., Gkritzalis, T., Goncalves De Souza, J., Grassi, G., Gregor, L., Gruber, N., Guenet, B., Gürses, O., Harrington, K., Harris, I., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Ito, A., Jacobson, A. R., Jain, A. K., Jarníková, T., Jersild, A., Jiang, F., Jones, S. D., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Kong, Y., Korsbakken, J. I., Koven, C., Kunimitsu, T., Lan, X., Liu, J., Liu, Z., Liu, Z., Lo Monaco, C., Ma, L., Marland, G., McGuire, P. C., McKinley, G. A., Melton, J. R., Monacci, N., Monier, E., Morgan, E. J., Munro, D. R., Müller, J. D., Nakaoka, S.-I., Nayagam, L. R., Niwa, Y., Nutzel, T., Olsen, A., Omar, A. M., Pan, N., Pandey, S., Pierrot, D., Qin, Z., Regnier, P., Rehder, G., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Skjelvan, I., Smallman, T. L., Spada, V., Sreeush, M. G., Sun, Q., Sutton, A. J., Sweeney, C., Swingedouw, D., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tian, X., Tilbrook, B., Tsujino, H., Tubiello, F., van Ooijen, E., van der Werf, G. R., van de Velde, S. J., Walker, A. P., Wanninkhof, R., Yang, X., Yuan, W., Yue, X., and Zeng, J. (2026). Global carbon budget 2025. *Earth System Science Data*, 18(5):3211–3288.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Landschützer, P., Le Quéré, C., Li, H., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Arneth, A., Arora, V., Bates, N. R., Becker, M., Bellouin, N., Berghoff, C. F., Bittig, H. C., Bopp, L., Cadule, P., Campbell, K., Chamberlain, M. A., Chandra, N., Chevallier, F., Chini, L. P., Colligan, T., Decayeux, J., Djeutchouang, L., Dou, X., Duran Rojas, C.,

Enyo, K., Evans, W., Fay, A., Feely, R. A., Ford, D. J., Foster, A., Gasser, T., Gehlen, M., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, O., Harris, I., Hefner, M., Heinke, J., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Kato, E., Keeling, R. F., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Lauvset, S. K., Lefèvre, N., Liu, Z., Liu, J., Ma, L., Maksyutov, S., Marland, G., Mayot, N., McGuire, P., Metzl, N., Monacci, N. M., Morgan, E. J., Nakaoka, S.-I., Neill, C., Niwa, Y., Nützel, T., Olivier, L., Ono, T., Palmer, P. I., Pierrot, D., Qin, Z., Resplandy, L., Roobaert, A., Rosan, T. M., Rödenbeck, C., Schwinger, J., Smallman, T. L., Smith, S., Sospedra-Alfonso, R., Steinhoff, T., Sun, Q., Sutton, A. J., Séférian, R., Takao, S., Tatebe, H., Tian, H., Tilbrook, B., Torres, O., Tourigny, E., Tsujino, H., Tubiello, F., van der Werf, G., Wanninkhof, R., Wang, X., Yang, D., Yang, X., Yu, Z., Yuan, W., Yue, X., Zaehle, S., Zeng, N., and Zeng, J. (2024). Global carbon budget 2024. *Earth System Science Data Discussions*, 2024:1–133.

Fritsch, F. N. and Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246.

Geurts, P., Ernst, D., and Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1):3–42.

Giglio, L., Randerson, J. T., and van der Werf, G. R. (2013). Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (gfed4). *Journal of Geophysical Research: Biogeosciences*, 118(1):317–328.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B., Randerson, J., Sarmiento, J. L., Taguchi, S., Takahashi, T., Tans, P., and Yuen, C.-W. (2002). Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. *Nature*, 415.

Hansen, P. C. (1998). *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM.

Haynes, K. D., Baker, I. T., Denning, A. S., Stöckli, R., Schaefer, K., Lokupitiya, E. Y., and Haynes, J. M. (2019a). Representing grasslands using dynamic prognostic phenology based on biological growth stages: 1. implementation in the Simple Biosphere model (SiB4). *Journal of Advances in Modeling Earth Systems*, 11(12):4423–4439.

Haynes, K. D., Baker, I. T., Denning, A. S., Wolf, S., Wohlfahrt, G., Kiely, G., Minaya, R. C., and Haynes, J. M. (2019b). Representing grasslands using dynamic prognostic phenology based on biological growth stages: Part 2. carbon cycling. *Journal of Advances in Modeling Earth Systems*, 11(12):4440–4465.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Houtekamer, P. L. and Mitchell, H. L. (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796 – 811.

Jacobson, A. R., Gruber, N., Sarmiento, J. L., Gloor, M., and Fletcher, S. E. M. (2007). A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: I. methods and global-scale fluxes. *Global Biogeochemical Cycles*, 21(GB1019).

Jacobson, A. R., Schuldt, K. N., Miller, J. B., Pera, A., Kaushik, A., Andrews, A., Basu, S., Triñanes, J., Landschützer, P., Mund, J., Weir, B., Ott, L., Aalto, T., Aaltonen, H., Abshire, J. B., Aikin, K., Allen, G., Andrade, M., Apadula, F., Arnold, S., Baier, B., Bakwin, P., Bartyzel, J., Bentz, G., Bergamaschi, P., Beyersdorf, A., Biermann, T., Biraud, S. C., Blanc, P.-E., Boenisch, H., Bowling, D., Brailsford, G., Brand, W. A., Brunner, D., Bui, T. P. V., Burban, B., Bäni, L., Calzolari, F., Chang, C. S., Chen, G., Chen, H., Chmura, L., Clair, J. M. S., Clark, S., Climadat, S., Coletta, J. D., Colomb, A., Commane, R., Condori, L., Conen, F., Conil, S., Couret, C., Cristofanelli, P., Cuevas, E., Curcoll, R., Daube, B., Davis, K. J., Mazière, M. D., de Souza, R. A. F., Wekker, S. D., Dean-Day, J. M., Delmotte, M., Iorio, T. D., di Sarra, A. G., Dickerson, R., DiGangi, E., DiGangi, J. P., Elsasser, M., Emmenegger, L., Fang, S., Fischer, M. L., Forster, G., France, J., Frumau, A., Fuente-Lastra, M., Galkowski, M., Gatti, L. V., Gehrlein, T., Gerbig, C., Gheusi, F., Gloor, E., Goto, D., Griffis, T., Hammer, S., Hanisco, T. F., Hanson, C., Hashimoto, S., Haszpra, L., Hatakka, J., Heimann, M., Heliasz, M., Heltai, D., Henne, S., Hensen, A., Hermans, C., Hermansen, O., Higgs, J., Hintsa, E., Hoheisel, A., Holst, J., Iraci, L. T., Ivakhov, V., Jaffe, D. A., Joly, L., Jordan, A., Joubert, W., Kang, H.-Y., Karion, A., Kawa, S. R., Kazan, V., Keeling, R. F., Kentaro, I., Keronen, P., Kim, J., Klausen, J., Kneuer, T., Ko, M.-Y., Kolari, P., Komínková, K., Kort, E., Kozlova, E., Krummel, P., Kubistin, D., Kulawik, S. S., Kumps, N., Labuschagne, C., Lam, D. H., Lan, X., Langenfelds, R. L., Lanza, A., Larmanou, E., Laurent, O., Lauvaux, T., Lavric, J., Law, B. E., Lee, C.-H., Lee, J., Lehner, I., Lehtinen, K., Leppert, R., Leskinen, A., Leuenberger, M., Leung, W., Levin, I., Levula, J., Lin, J., Lindauer, M., Lindroth, A., Loh, Z., Lopez, M., Lueker, T. J., Luijkx, I. T., Lunder, C. R., Machida, T., Mammarella, I., Manca, G., Manning, A., Manning, A., Marek, M. V., Marklund, P., Marrero, J. E., Martin, D., Martin, M. Y., Martins, G. A., Matsueda, H., McAuliffe, A., McKain, K., Meijer, H., Meinhardt, F., Merchant, L., Metzger, J.-M., Mihalopoulos, N., Miles, N. L., Miller, C. E., Mitchell, L., Mölder, M., Müller-Williams, J., Monteiro, V., Montzka, S., Moossen, H., Moreno, C., Morgan, E., Morgui, J.-A., Morimoto, S., Mukai, H., Munger, J. W., Munro, D., Mutuku, M., Myhre, C. L., Nakaoka, S.-I., Necki, J., Newberger, T., Newman, S., Nichol, S., Nisbet, E., Niwa, Y., Njiru, D. M., Noe, S. M., Nojiri, Y., Obersteiner, F., O'Doherty, S., Paplawsky, B., Parworth, C. L., Peischl, J., Peltola, O., Peters, W., Philippon, C., Piacentino, S., Pichon, J.-M., Pickers, P., Piper, S., Pitt, J., Plass-Dülmer, C., Platt, S. M., Prinzivalli, S., Ramonet, M., Ramos, R., Ren, X., Reyes-Sanchez, E., Richardson, S. J., Rigouleau, L.-J., Riris, H., Rivas, P. P., Rothe, M., Roulet, Y.-A., Ryerson, T., Ryoo, J.-M., Sargent, M., Sasakawa, M., Scheeren, B., Schmidt, M., Schuck, T., Schumacher, M., Seibel, J., Seifert, T., Sha, M. K., Shepson, P., Shin, D., Shook, M., Sloop, C. D., Smale, D., Spain, G., Stavert, A., Steger, D., Steinbacher, M., Stephens, B., Sweeney, C., Sørensen, L. L., Taipale, R., Takatsuji, S., Tans, P., Terao, Y., Thoning, K., Timas, H., Torn, M., Trisolino, P., Tørseth, K., Turnbull, J., van den Bulk, P., Vermeulen, A., Viner, B., Vitkova, G., Walker, S., Watson, A., Weiss, R., Weyrauch, D., Wofsy, S. C., Wolter, S., Worsey, J., Worthy, D., Xueref-Remy, I., Yates, E. L., Young, D., Yver-Kwok, C., Zaehle, S., Zahn, A., Zellweger, C., and Zimnoch, M. (2026). CarbonTracker CT2026.

Jersild, A., Landschützer, P., Gruber, N., and Bakker, D. C. E. (2017). An observation-based global monthly gridded sea surface pCO_2 and air-sea CO_2 flux product from 1982 onward and its monthly climatology. Dataset available at NCEI https://doi.org/10.7289/v5z899n6.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G. (2001). On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research-Atmospheres*, 106(D5):4703–4715.

Kohonen, T. (1987). *Self-Organization and Associative Memory*, volume 8 of *Springer Series in Information Sciences*. Springer-Verlag, Berlin, Heidelberg.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P. (2005). The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chemistry and Physics*, 5:417–432.

Landschützer, P., Gruber, N., Bakker, D., Schuster, U., Nakaoka, S., Payne, M., Sasse, T., and Zeng, J. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, 10(11):7793.

Landschützer, P., Gruber, N., and Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, 30(10):1396–1417.

Levitus, S., Locarnini, R. A., Boyer, T. P., Mishonov, A. V., Antonov, J. I., Garcia, H. E., Baranova, O. K., Zweng, M. M., Johnson, D. R., and Seidov, D. (2010). *World Ocean Atlas 2009*. NOAA National Centers for Environmental Information.

Liu, Z., Ciais, P., Deng, Z., Davis, S. J., Zheng, B., Wang, Y., Cui, D., Zhu, B., Dou, X., Ke, P., Sun, T., Guo, R., Zhong, H., Boucher, O., Bréon, F.-M., Lu, C., Guo, R., Xue, J., Boucher, E., Tanaka, K., and Chevallier, F. (2020). Carbon Monitor, a near-real-time daily dataset of global CO_2 emission from fossil fuel and cement production. *Scientific Data*, 7(1).

Lokupitiya, E., Denning, A. S., Schaefer, K., Ricciuto, D., Anderson, R., Arain, M. A., Baker, I., Barr, A. G., Chen, G., Chen, J. M., Ciais, P., Cook, D. R., Dietze, M., El Maayar, M., Fischer, M., Grant, R., Hollinger, D., Izaurralde, C., Jain, A., Kucharik, C., Li, Z., Liu, S., Li, L., Matamala, R., Peylin, P., Price, D., Running, S. W., Sahoo, A., Sprintsin, M., Suyker, A. E., Tian, H., Tonitto, C., Torn, M., Verbeeck, H., Verma, S. B., and Xue, Y. (2016). Carbon and energy fluxes in cropland ecosystems: a model-data comparison. *Biogeochemistry*, 129(1):53–76.

Marland, G. (2008). Uncertainties in accounting for CO_2 from fossil fuels. *Journal of Industrial Ecology*, 12(2):136–139.

Mu, M., Randerson, J. T., van der Werf, G. R., Giglio, L., Kasibhatla, P., Morton, D., Collatz, G. J., DeFr̃ies, R. S., Hyer, E. J., Prins, E. M., Griffith, D. W. T., Wunch, D., Toon, G. C., Sherlock, V., and Wennberg, P. O. (2011). Daily and 3-hourly variability in global fire emissions and consequences for atmospheric model predictions of carbon monoxide. *Journal of Geophysical Research: Atmospheres*, 116(D24).

Nassar, R., Napier-Linton, L., Gurney, K. R., Andres, R. J., Oda, T., Vogel, F. R., and Deng, F. (2013). Improving the temporal and spatial distribution of CO_2 emissions from global fossil fuel emission data sets. *Journal of Geophysical Research: Atmospheres*, 118(2):917–933.

Olsen, S. C. and Randerson, J. T. (2004). Differences between surface and column atmospheric CO_2 and implications for carbon cycle research. *Journal of Geophysical Research-Atmospheres*, 109(D2).

Olson, J., Watts, J., and Allsion, L. (1992). Major world ecosystem complexes ranked by carbon in live vegetation: A database. Technical Report ORNL/CDIAC-134, NDP-017, Carbon Dioxide Information Analysis Center, U.S. Department of Energy, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

Patra, P. K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M. P., Corbin, K., et al. (2011). TransCom model simulations of CH_4 and related

species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere. *Atmospheric Chemistry and Physics*, 11(24):12813–12837.

Pera, A., Jacobson, A. R., and Miller, J. B. (2026). Gridded $1^\circ \times 1^\circ$ monthly fossil-fuel CO_2 emission estimates (1993–2025) for use as prior fluxes in atmospheric inversion systems. Data product version 2026; pipeline code at https://github.com/noaa-gml/Miller-Pera-FF.

Peters, W., Krol, M. C., Dlugokencky, E. J., Dentener, F. J., Bergamaschi, P., Dutton, G., von Velthoven, P., Miller, J. B., Bruhwiler, L., and Tans, P. P. (2004). Toward regional-scale modeling using the two-way nested global model TM5: Characterization of transport using SF_6. *Journal of Geophysical Research-Atmospheres*, 109(D19). D19314.

Peters, W., Miller, J., Whitaker, J., Denning, A., Hirsch, A., Krol, M., Zupanski, D., Bruhwiler, L., and Tans, P. (2005). An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations. *Journal of Geophysical Research-Atmospheres*, 110:D24304. D24304.

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., and Klooster, S. A. (1993). Terrestrial ecosystem production - a process model-based on global satellite and surface data. *Global Biogeochemical Cycles*, 7(4):811–841.

Randerson, J., van der Werf, G., Giglio, L., Collatz, G., and Kasibhatla, P. (2017). Global Fire Emissions Database, version 4.1 (GFEDv4).

Randerson, J. T., Chen, Y., van der Werf, G. R., Rogers, B. M., and Morton, D. C. (2012). Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences*, 117(G4).

Rasmussen, L. (1991). Piecewise integral splines of low degree. *Computers & Geosciences*, 17(9):1255–1263.

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., Wanninkhof, R., Wong, C. S., Wallace, D. W. R., Tilbrook, B., Millero, F. J., Peng, T. H., Kozyr, A., Ono, T., and Rios, A. F. (2004). The oceanic sink for anthropogenic CO_2. *Science*, 305(5682):367–371.

Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., Prihodko, Lara an d Suits, N., and Philpott, A. (2008). Combined Simple Biosphere/Carnegie-Ames-Stanford Approach terrestrial carbon cycle model. *Journal of Geophysical Research: Biogeosciences*, 113(G3).

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., Chevallier, F., Crowell, S., Davis, K. J., Deng, F., et al. (2019). Quantifying the impact of atmospheric transport uncertainty on CO_2 surface flux estimates. *Global biogeochemical cycles*, 33(4):484–500.

Schuldt, K. N., Mund, J., Aalto, T., Aaltonen, H., Abshire, J. B., Aikin, K., Allen, G., Andrade, M., Apadula, F., Arnold, S., Baier, B., Bakwin, P., Bäni, L., Bartyzel, J., Bentz, G., Bergamaschi, P., Beyersdorf, A., Biermann, T., Biraud, S. C., Blanc, P.-E., Boenisch, H., Bowling, D., Brailsford, G., Brand, W. A., Brunner, D., Bui, T. P. V., van den Bulk, P., Burban, B., Calzolari, F., Chang, C. S., Chen, G., Chen, H., Chmura, L., St. Clair, J. M., Clark, S., Climadat, S., Coletta, J. D., Colomb, A., Commane, R., Condori, L., Conen, F., Conil, S., Couret, C., Cristofanelli, P., Cuevas, E., Curcoll, R., Daube, B., Davis, K. J., Dean-Day, J. M., Delmotte, M., Dickerson, R., DiGangi, E., DiGangi, J. P., Elsasser, M., Emmenegger, L., Fang, S., Forster, G., France, J., Frumau, A., Fuente-Lastra, M., Galkowski, M., Gatti, L. V., Gehrlein, T., Gerbig, C., Gheusi, F., Gloor, E., Goto, D., Griffis, T., Hammer, S., Hanisco, T. F., Hanson, C., Hashimoto, S., Haszpra, L., Hatakka, J., Heimann, M., Heliasz, M., Heltai, D., Henne, S., Hensen, A., Hermans, C.,

Hermansen, O., Higgs, J., Hintsa, E., Hoheisel, A., Holst, J., Di Iorio, T., Iraci, L. T., Ivakhov, V., Jaffe, D. A., Joly, L., Jordan, A., Joubert, W., Kang, H.-Y., Karion, A., Kawa, S. R., Kazan, V., Keeling, R. F., Kentaro, I., Keronen, P., Kim, J., Klausen, J., Kneuer, T., Ko, M.-Y., Kolari, P., Kominkova, K., Kort, E., Kozlova, E., Krummel, P. B., Kubistin, D., Kulawik, S. S., Kumps, N., Labuschagne, C., Lam, D. H. Y., Lan, X., Langenfelds, R. L., Lanza, A., Larmanou, E., Laurent, O., Lauvaux, T., Lavric, J., Law, B. E., Lee, C.-H., Lee, J., Lehner, I., Lehtinen, K., Leppert, R., Leskinen, A., Leuenberger, M., Leung, W. H., Levin, I., Levula, J., Lin, J., Lindauer, M., Lindroth, A., Loh, Z. M., Lopez, M., Lueker, T. J., Lunder, C. R., Machida, T., Mammarella, I., Manca, G., Manning, A., Manning, A., Marek, M. V., Marklund, P., Marrero, J. E., Martin, D., Martin, M. Y., Martins, G. A., Matsueda, H., De Mazière, M., McAuliffe, A., McKain, K., Meijer, H., Meinhardt, F., Merchant, L., Metzger, J.-M., Mihalopoulos, N., Miles, N. L., Miller, C. E., Miller, J. B., Mitchell, L., Mölder, M., Monteiro, V., Montzka, S., Moossen, H., Moreno, C., Morgan, E., Morgui, J.-A., Morimoto, S., Mukai, H., Müller-Williams, J., Munger, J. W., Munro, D., Mutuku, M., Myhre, C. L., Nakaoka, S.-I., Necki, J., Newberger, T., Newman, S., Nichol, S., Nisbet, E., Niwa, Y., Njiru, D. M., Noe, S. M., Nojiri, Y., O'Doherty, S., Obersteiner, F., Paplawsky, B., Parworth, C. L., Peischl, J., Peltola, O., Peters, W., Philippon, C., Piacentino, S., Pichon, J. M., Pickers, P., Piper, S., Pitt, J., Plass-Dülmer, C., Platt, S. M., Prinzivalli, S., Ramonet, M., Ramos, R., Ren, X., Reyes-Sanchez, E., Richardson, S. J., Rigouleau, L.-J., Riris, H., Rivas, P. P., Rothe, M., Roulet, Y.-A., Ryerson, T., Ryoo, J.-M., Sargent, M., di Sarra, A. G., Sasakawa, M., Scheeren, B., Schmidt, M., Schuck, T., Schumacher, M., Seibel, J., Seifert, T., Sha, M. K., Shepson, P., Shin, D., Shook, M., Sloop, C. D., Smale, D., Sørensen, L. L., de Souza, R. A. F., Spain, G., Stavert, A., Steger, D., Steinbacher, M., Stephens, B., Sweeney, C., Taipale, R., Takatsuji, S., Tans, P., Terao, Y., Thoning, K., Timas, H., Torn, M., Trisolino, P., Turnbull, J., Vermeulen, A., Viner, B., Vitkova, G., Walker, S., Watson, A., Weiss, R., De Wekker, S., Weyrauch, D., Wofsy, S. C., Wolter, S., Worsey, J., Worthy, D., Xueref-Remy, I., Yates, E. L., Young, D., Yver-Kwok, C., Zaehle, S., Zahn, A., Zellweger, C., and Zimnoch, M. (2026). Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2024; obspack_co2_1_globalviewplus_v11.0_2026-01-13.

Stöckli, R., Rutishauser, T., Baker, I., Liniger, M. A., and Denning, A. S. (2011). A global reanalysis of vegetation phenology. *Journal of Geophysical Research: Biogeosciences*, 116(G3).

Takahashi, T., Sutherland, S. C., Sweeney, C., Metzl, A. P. N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R. A., Sabine, C., Olafsson, J., and Nojiri, Y. (2002). Global air-sea CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep-Sea Research II*, 49:1601–1622.

Thoning, K. W., Tans, P. P., and Komhyr, W. D. (1989). Atmospheric carbon dioxide at Mauna Loa observatory. 2. Analysis of the NOAA GMCC data, 1974-1985. *Journal of Geophysical Research-Atmospheres*, 94:8549–8565.

Tomohiro Oda, S. M. (2015). ODIAC fossil fuel CO_2 emissions dataset (ODIAC2024).

van der Werf, G., Randerson, J., Collatz, G., Giglio, L., Kasibhatla, P., Arellano, A., Olsen, S., and Kasischke, E. (2004). Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period. *Science*, 303:73–76.

van der Werf, G. R., Randerson, J. T., van Wees, D., Chen, Y., Giglio, L., Hall, J., Vernooij R., Mu, M., Binte Shahid, S., Barsanti, K. C., Yokelson, R., and Morton, D. C. (2025). Landscape fire emissions from the 5th version of the Global Fire Emissions Database (GFED5). *Scientific Data*, 12(1):1870.

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, 12(6):351–362.

Wanninkhof, R., Triñanes, J., Pierrot, D., Munro, D. R., Sweeney, C., and Fay, A. R. (2025). Trends in sea-air co2 fluxes and sensitivities to atmospheric forcing using an extremely randomized trees machine learning approach. *Global Biogeochemical Cycles*, 39(2):e2024GB008315. e2024GB008315 2024GB008315.

Wanninkhof, R., Triñanes, J. A., Pierrot, D., Munro, D. R., Sweeney, C., and Fay, A. R. (2024). AOML_ET: Partial pressure of CO_2 (pCO_2) and sea-air CO_2 fluxes for the global ocean, along with the predictor variables from 19 98-01-01 to 2023-12-30, using an Extra Trees (extremely randomized trees) machine learning. Dataset available at NCEI https://doi.org/10.25921/0s8y-q287.

Weir, B., Ott, L. E., Collatz, G. J., Kawa, S. R., Poulter, B., Chatterjee, A., Oda, T., and Pawson, S. (2021). Bias-correcting carbon fluxes derived from land-surface satellite data for retrospective and near-real-time assimilation systems. *Atmospheric Chemistry and Physics*, 21(12):9609–9628.

Whitaker, J. S. and Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7):1913–1924.

Appendix A

Full author list

Andrew R. Jacobson[1,2], Kenneth N. Schuldt[1,2], John B. Miller[2], Ashley Pera[1,2], Aleya Kaushik[1,2], Arlyn Andrews[2(retired),3], Sourish Basu[4,5], Joaquin Triñanes[6,7,8], Peter Landschützer[9], Brad Weir[10,4], Lesley Ott[4], John Mund[1,2], Tuula Aalto[11], Hermanni Aaltonen[11], James Brice Abshire[12], Ken Aikin[13], Grant Allen[14], Marcos Andrade[15], Francesco Apadula[16], Sabrina Arnold[17], Bianca Baier[2], Peter Bakwin, Jakub Bartyzel[18], Gilles Bentz[19], Peter Bergamaschi[20], Andreas Beyersdorf[21], Tobias Biermann[22], Sebastien C. Biraud[23], Pierre-Eric Blanc[24], Harald Boenisch[25], David Bowling[26], Gordon Brailsford[27], Willi A. Brand[28], Dominik Brunner[29], Thao Paul V. Bui[30], Benoit Burban[31], Lukas Bäni[32], Francescopiero Calzolari[33], Cecilia S. Chang[30], Gao Chen[34], Huilin Chen[35], Lukasz Chmura[18], Jason M. St. Clair[12], Shane Clark[36], Sites Climadat[37], Julian Della Coletta[38], Aurelie Colomb[39], Roisin Commane[40], Lino Condori[41], Franz Conen[42], Sébastien Conil[43], Cédric Couret[44], Paolo Cristofanelli[33], Emilio Cuevas[45], Roger Curcoll[37], Bruce Daube[40], Kenneth J. Davis[46], Martine De Mazière[47], Rodrigo A. F. de Souza[48], Stephan De Wekker[49], Jonathan M. Dean-Day[30], Marc Delmotte[19], Tatiana Di Iorio[50], Alcide Giorgio di Sarra[50], Russell Dickerson[15], Elizabeth DiGangi[51], Joshua P. DiGangi[34], Michael Elsasser[52], Lukas Emmenegger[29], Shuangxi Fang[53], Marc L. Fischer[54], Grant Forster[55,56], James France[57], Arnoud Frumau[58], Marta Fuente-Lastra[19], Michal Galkowski[18], Luciana V. Gatti[59], Torsten Gehrlein[25], Christoph Gerbig[28], Francois Gheusi[60], Emanuel Gloor[61], Daisuke Goto[62], Tim Griffis[63], Samuel Hammer[38], Thomas F. Hanisco[12], Chad Hanson[64], Shigeru Hashimoto[65], László Haszpra[66], Juha Hatakka[11], Martin Heimann[28], Michal Heliasz[22], Daniela Heltai[16], Stephan Henne[29], Arjan Hensen[58], Christian Hermans[67], Ove Hermansen[68], Jack Higgs[2], Eric Hintsa[2], Antje Hoheisel[38], Jutta Holst[69], Laura T. Iraci[30], Viktor Ivakhov[70], Daniel A. Jaffe[71], Lilian Joly[72], Armin Jordan[28], Warren Joubert[73], Hui-Yun Kang[74], Anna Karion[75], Stephan Randolph Kawa[12], Victor Kazan[19], Ralph F. Keeling[36], Ishijima Kentaro[65], Petri Keronen[76], Jooil Kim[36], Jörg Klausen[77], Tobias Kneuer[78], Mi-Young Ko[74], Pasi Kolari[76], Kateřina Komínková[79], Eric Kort[80], Elena Kozlova[81], Paul Krummel[82], Dagmar Kubistin[78], Susan S. Kulawik[30], Nicolas Kumps[67], Casper Labuschagne[73], David H.Y. Lam[83], Xin Lan[1,2], Ray L. Langenfelds[82], Andrea Lanza[16], Eric Larmanou[84], Olivier Laurent[85],

Thomas Lauvaux[72,46], Jost Lavric[28], Beverly E. Law[64], Choong-Hoon Lee[74],
John Lee[86], Irene Lehner[22], Kari Lehtinen[11], Reimo Leppert[28],
Ari Leskinen[87,88], Markus Leuenberger[32], W.H. Leung[83], Ingeborg Levin[38],
Janne Levula[76], John Lin[26], Matthias Lindauer[78], Anders Lindroth[69],
Zoe Loh[82], Morgan Lopez[19], Timothy J. Lueker[36], Ingrid T. Luijkx[89,90],
Chris René Lunder[68], Toshinobu Machida[65], Ivan Mammarella[91], Giovanni Manca[20],
Alistair Manning[92], Andrew Manning[56], Michal V. Marek[79], Per Marklund[84],
Josette E. Marrero[30], Damien Martin[93], Melissa Yang Martin[34], Giordane A. Martins[94],
Hidekazu Matsueda[95], Anna McAuliffe[1,2], Kathryn McKain[2], Harro Meijer[35],
Frank Meinhardt[96], Lynne Merchant[36], Jean-Marc Metzger[47], N. Mihalopoulos[97],
Natasha L. Miles[46], Charles E. Miller[98], Logan Mitchell[26], Meelis Mölder[69],
Jennifer Müller-Williams[78], Vanessa Monteiro[46], Stephen Montzka[2], Heiko Moossen[28],
Caisa Moreno, Eric Morgan[36], Josep-Anton Morgui[37], Shinji Morimoto[99],
Hitoshi Mukai[65], J. William Munger[40], David Munro[2], Mathew Mutuku[100],
Cathrine Lund Myhre[68], Shin-Ichiro Nakaoka[65], Jaroslaw Necki[18], Tim Newberger[2],
Sally Newman[101], Sylvia Nichol[27], Euan Nisbet[57], Yosuke Niwa[65],
David Murithi Njiru[100], Steffen Manfred Noe[102], Yukihiro Nojiri[65], Simon O'Doherty[103],
Florian Obersteiner[25], Simon O'Doherty[103], Bill Paplawsky[36], Caroline L. Parworth[30],
Jeff Peischl[1,13], Olli Peltola[76], Wouter Peters[35], Carole Philippon[19],
Salvatore Piacentino[50], Jean-Marc Pichon[39], Penelope Pickers[56], Steve Piper[36],
Joseph Pitt[103], Christian Plass-Dülmer[78], Stephen Matthew Platt[68], Steve Prinzivalli[51],
Michel Ramonet[19], Ramon Ramos[45], Xinrong Ren[104], Enrique Reyes-Sanchez[45],
Scott J. Richardson[46], Louis-Jeremy Rigouleau[19], Haris Riris[12], Pedro P. Rivas[45],
Michael Rothe[28], Yves-Alain Roulet[77], Thomas Ryerson[13], Ju-Mee Ryoo[30],
Maryann Sargent[40], Motoki Sasakawa[65], Bert Scheeren[35], Martina Schmidt[19],
Tanja Schuck[105], Marcus Schumacher[28], Jennifer Seibel[36], Thomas Seifert[28],
Mahesh Kumar Sha[67], Paul Shepson[106], Daegeun Shin[74], Michael Shook[34],
Christopher D. Sloop[51], Dan Smale[27], Gerard Spain[93], Ann Stavert[82],
David Steger[29], Martin Steinbacher[29], Britton Stephens[107], Colm Sweeney[2],
Lise Lotte Sørensen[108], Risto Taipale[76], Shinya Takatsuji[109], Pieter Tans[110],
Yukio Terao[65], Kirk Thoning[2], Helder Timas[111], Margaret Torn[54],
Pamela Trisolino[33], Kjetil Tørseth[68], Jocelyn Turnbull[112,1], Pim van den Bulk[58],
Alex Vermeulen[69], Brian Viner[113], Gabriela Vitkova[79], Stephen Walker[36],
Andrew Watson[81], Ray Weiss[36], Dietmar Weyrauch[78], Steven C. Wofsy[40],
Sonja Wolter[2], Justin Worsey[81], Doug Worthy[114], Irène Xueref-Remy[115],
Emma L. Yates[30], Dickon Young[103], Camille Yver-Kwok[19], Sönke Zaehle[28],
Andreas Zahn[25], Christoph Zellweger[29] and Miroslaw Zimnoch[18]

[1]*CIRES, University of Colorado, Boulder, Colorado, USA*
[2]*NOAA Global Monitoring Laboratory, Boulder, Colorado, USA*
[3]*Former Feds and Friends, LLC*
[4]*Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[5]*Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA*
[6]*Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, FL*
[7]*Rosenstiel School of Marine, Atmospheric and Earth Science, Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL*
[8]*Department of Electronics and Computer Science, Universidadede Santiago de Compostela, Santiago, Spain*

[9]*Vlaams Instituut voor de Zee, Oostende, België*
[10]*Morgan State University, Baltimore, Maryland, USA*
[11]*Finnish Meteorological Institute, Climate System Research, Helsinki, Finland*
[12]*NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[13]*NOAA Chemical Sciences Laboratory, Boulder, Colorado, USA*
[14]*University of Manchester, Manchester, England*
[15]*Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA*
[16]*Ricerca sul Sistema Energetico– RSE S.p.A., Milano, Italy*
[17]*Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany*
[18]*AGH University of Science and Technology, Krakow, Poland*
[19]*Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, Gif-sur-Yvette, France*
[20]*European Commission, Joint Research Centre, Ispra, Italy*
[21]*California State University, San Bernardino, California, USA*
[22]*Lund University, Centre for Environmental and Climate Science, Lund, Sweden*
[23]*ARM Carbon Project, Lawrence Berkeley National Laboratory, Berkeley, California, USA*
[24]*Observatory of Universe Sciences Pytheas, Aix-Marseille University, Marseille, France*
[25]*Institute for Meteorology and Climate Research (IMK), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*
[26]*Department of Atmospheric Sciences, University of Utah, Salt Lake City, Utah, USA*
[27]*National Institute of Water and Atmospheric Research, Wellington, New Zealand*
[28]*Max Planck Institute for Biogeochemistry, Jena, Germany*
[29]*Empa, Swiss Federal Laboratories for Materials Science and Technology, Laboratory for Air Pollution/Environmental Technology, Dübendorf, Switzerland*
[30]*NASA Ames Research Center, Moffett Field, California, USA*
[31]*Joint Research Unit Ecology of Guianan Forests*
[32]*Climate and Environmental Physics, University of Bern, Bern, Switzerland*
[33]*Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, Italy*
[34]*NASA Langley Research Center, Hampton, Virginia, USA*
[35]*Centre for Isotope Research, University of Groningen, Groningen, Netherlands*
[36]*Scripps Institution of Oceanography, University of California, La Jolla, California, USA*
[37]*Institut de Ciencia i Tecnologia Ambientals, Universitat Autonoma de Barcelona, Barcelona, Spain*
[38]*Universität Heidelberg, Institut für Umweltphysik, Heidelberg, Germany*
[39]*Observatoire de Physique du Globe de Clermont Ferrand, Aubiere, France*
[40]*Harvard University, School of Engineering and Applied Sciences, Cambridge, Massachusetts, USA*
[41]*Servicio Meteorológico Nacional, Argentina*
[42]*University of Basel, Basel, Switzerland*
[43]*Agence Nationale pour la Gestion des Déchets Radioactifs, France*
[44]*Umweltbundesamt, Zugspitze, Germany*
[45]*Agencia Estatal Meteorologia, Santa Cruz de Tenerife, Spain*
[46]*The Pennsylvania State University, Department of Meteorology and Atmospheric Science, University Park, Pennsylvania, USA*
[47]*Observatoire de Physique de l'Atmosphère de la Réunion*
[48]*Amazonas State University, Manaus, Brazil*
[49]*University of Virginia, Charlottesville, Virginia, USA*
[50]*Italian National Agency for New Technologies, Energy and Sustainable Economic Development, UTMEA-TER Earth Observations and Analyses Laboratory, Rome, Italy*
[51]*Earth Networks, Inc., an AEM company, Germantown, Maryland, USA*
[52]*Umweltbundesamt, Salzburg, Germany*
[53]*Meteorological Observation Centre, Chinese Meteorological Administration, Beijing, China*
[54]*Lawrence Berkeley National Laboratory, Berkeley, California, USA*
[55]*National Centre for Atmospheric Sciences, University of East Anglia, Norwich, Norfolk, United Kingdom*
[56]*Centre for Ocean and Atmospheric Sciences, University of East Anglia, Norfolk, United Kingdom*
[57]*Royal Holloway, University of London, United Kingdom*
[58]*Netherlands Organisation for Applied Scientific Research (TNO), Petten, The Netherlands*

[59]*National Institute for Space Research (INPE), Sao Paulo, Brazil*
[60]*Observatoire Midi-Pyrénées, Toulouse, France*
[61]*University of Leeds,School of Geography, Leeds, United Kingdom*
[62]*National Institute of Polar Research, Tokyo, Japan*
[63]*University of Minnesota,Department of Soil, Water, and Climate, St. Paul, Minnesota, USA*
[64]*Oregon State University, Corvallis, Oregon, USA*
[65]*National Instiute for Environmental Studies, Tsukuba, Japan*
[66]*Institute for Nuclear Research, Debrecen, Hungary*
[67]*Royal Belgian Institute for Space Aeronomy, Brussels, Belgium*
[68]*NILU, Kjeller, Norway*
[69]*Lund University, Dept. Phys. Geography and Ecosystem Science, Lund, Sweden*
[70]*Voeikov Main Geophysical Observatory,Saint Petersburg, Russia*
[71]*University of Washington, Seattle, Washington, USA*
[72]*University of Reims Champagne-Ardenne, CNRS, Reims, France*
[73]*South African Weather Service, Cape Point, South Africa*
[74]*Korea Meteorological Administration, Seoul, South Korea*
[75]*National Institute of Standards and Technology, Gaithersburg, Maryland, USA*
[76]*University of Helsinki, Helsinki, Finland*
[77]*Federal Office of Meteorology and Climatology MeteoSwiss, Switzerland*
[78]*Deutscher Wetterdienst, Hohenpeißenberg Meteorological Observatory, Hohenpeißenberg, Germany*
[79]*Global Change Research Institute of the Czech Academy of Sciences, Brno, Czech Republic*
[80]*University of Michigan, Ann Arbor, Michigan, USA*
[81]*University of Exeter, Centre for Environmental Data Analysis, Exeter, Devon, United Kingdom*
[82]*Commonwealth Scientific and Industrial Research Organisation, Environment, Aspendale, Victoria, Australia*
[83]*Hong Kong Observatory, Hong Kong, China*
[84]*Svartberget Field Research Station, Swedish University of Agricultural Sciences, Vindeln, Sweden*
[85]*ICOS Atmospheric Thematic Centre, Gif-sur-Yvette, France*
[86]*University of Maine, Orono, Maine, USA*
[87]*University of Eastern Finland, Department of Technical Physics, Kuopio, Finland*
[88]*Finnish Meteorological Institute, Kuopio, Finland*
[89]*Wageningen University, Wageningen, Netherlands*
[90] *ICOS Carbon Portal, Lund University, Lund, Sweden*
[91]*Institute for Atmospheric and Earth System Research/Physics, Faculty of Sciences, University of Helsinki, Finland*
[92]*Met Office Exeter, Devon, United Kingdom*
[93]*National University of Ireland, Galway, Ireland*
[94]*Fundacão Amazônica de Defesa da Atmosfera, Manaus, Brazil*
[95]*Meteorological Research Institute, Tsukuba, Japan*
[96]*Umweltbundesamt, Oberried-Hofsgrund, Germany*
[97]*Environmental and Chemical Processes Laboratory, University of Crete, Crete, Greece*
[98]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena California, USA*
[99]*Tohoku University, Sendai, Japan*
[100]*Kenya Meteorological Department, Nairobi, Kenya*
[101]*California Institute of Technology, Pasadena, California, USA*
[102]*Estonian University of Life Sciences, Institute of Forestry and Engineering, Tartu, Estonia*
[103]*University of Bristol, Bristol, United Kingdom*
[104]*NOAA Air Resources Laboratory, College Park, Maryland, USA*
[105]*Institute for Atmospheric and Environmental Sciences, University of Frankfurt, Frankfurt, Germany*
[106]*Purdue University, West Lafayette, Indiana, USA*
[107]*National Center for Atmospheric Research, Boulder, Colorado, USA*
[108]*Aarhus University, Aarhus Centrum, Denmark*
[109]*Japan Meteorological Agency, Tokyo, Japan*
[110]*Institute of Arctic and Alpine Research,University of Colorado, Boulder, Colorado, USA*
[111]*Instituto Nacional de Meteorologia e Geofisica, Cidade de Espargos, Ilha do Sal, República de Cabo Verde*
[112]*GNS Science,National Isotope Centre, Lower Hutt, New Zealand*

[113]*Savannah River National Laboratory, Aiken, South Carolina, USA*
[114]*Environment and Climate Change Canado, Ontario, Canada*
[115]*Aix-Marseille University, Marseille, France*

Appendix B

Performance by dataset

Table B.1 summarizes the datasets assimilated in CarbonTracker, and the performance of the assimilation scheme for each dataset. These diagnostics are useful for evaluating how well CarbonTracker does in simulating observed CO_2.

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_abp_surface-flask_1_representative	NOAA	Arembepe, Bahia, Brazil	12.77°S	38.17°W	1	91	0	0	0.6–2.1	0.46	-0.23	0.70
co2_abp_surface-flask_26_marine	IPEN	Arembepe, Bahia, Brazil	12.77°S	38.17°W	1	94	0	0	0.5–54.5	0.64	-1.97	12.65
co2_acg_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	423	557	33	0	0.2–7.3	0.87	-0.47	2.41
co2_acg_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	423	557	33	0	0.2–7.3	0.87	-0.47	2.41
co2_acg_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	1480	233	20	0	0.2–3.3	1.28	-0.47	1.48
co2_acg_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	2520	164	19	0	0.2–3.6	1.43	-0.30	1.47
co2_acg_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	3534	122	19	0	0.3–2.3	1.36	-0.19	1.35
co2_acg_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	4481	99	18	0	0.3–2.3	1.48	-0.10	1.28
co2_acg_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	5512	119	9	0	0.4–3.4	1.09	-0.00	1.11
co2_acg_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	6430	114	14	0	0.2–2.5	1.62	0.26	1.25
co2_acg_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	7485	141	13	0	0.1–2.9	1.16	0.09	1.34
co2_acg_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Alaska Coast Guard, United States	57.74°N	152.50°W	8320	13	0	0	0.4–1.8	0.17	-0.04	0.32

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_ah2_shipboard-insitu_20_allvalid	NIES	Alligator Hope (M/S Alligator Hope of Mitsui O.S.K. Lines, Ltd.)	variable		Surface	4609	493	0	0.2–13.2	1.29	-0.24	2.84
co2_alt_surface-flask_1_representative	NOAA	Alert, Nunavut, Canada	82.45°N	62.51°W	185	1606	54	6	0.3–4.2	0.83	-0.11	0.63
co2_alt_surface-flask_2_representative	CSIRO	Alert, Nunavut, Canada	82.45°N	62.51°W	185	939	64	0	0.1–2.9	1.18	-0.01	0.64
co2_alt_surface-flask_4_representative	SIO	Alert, Nunavut, Canada	82.45°N	62.51°W	185	591	12	1	0.3–4.1	1.06	-0.08	1.09
co2_alt_surface-flask_426_representative	SIO_CO2	Alert, Nunavut, Canada	82.45°N	62.51°W	185	471	17	0	0.3–3.1	1.07	-0.17	0.77
co2_alt_surface-insitu_6_allvalid	ECCC	Alert, Nunavut, Canada	82.45°N	62.51°W	185	29250	1691	0	0.4–2.2	1.21	-0.11	0.63
co2_ams_surface-insitu_11_allvalid	LSCE	Amsterdam Island, France	37.80°S	77.54°E	55	99111	1607	0	0.3–0.9	0.88	-0.10	0.58
co2_amt_surface-pfp_1_allvalid-107magl	NOAA	Argyle, Maine, United States	45.03°N	68.68°W	52	1751	9	0	1.1–13.0	0.46	0.09	2.40
co2_amt_tower-insitu_1_allvalid-107magl	NOAA	Argyle, Maine, United States	45.03°N	68.68°W	52	18696	328	0	1.1–7.9	0.64	0.07	3.03
co2_amt_tower-insitu_1_allvalid-12magl	NOAA	Argyle, Maine, United States	45.03°N	68.68°W	52	22315	407	0	1.2–5.5	0.68	0.15	3.68
co2_amt_tower-insitu_1_allvalid-30magl	NOAA	Argyle, Maine, United States	45.03°N	68.68°W	52	17582	326	0	1.2–5.2	0.69	0.44	3.35
co2_ara_surface-flask_2_representative	CSIRO	Arcturus, Queensland, Australia	23.86°S	148.47°E	175	16	0	0	1.0–4.7	0.41	0.17	1.89
co2_asc_surface-flask_1_representative	NOAA	Ascension Island, United Kingdom	7.97°S	14.40°W	85	1966	0	12	0.3–1.2	1.39	0.06	0.69
co2_ask_surface-flask_1_representative	NOAA	Assekrem, Algeria	23.26°N	5.63°E	2710	942	36	0	0.3–2.3	1.52	-0.22	0.86

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_azr_surface-flask_1_representative	NOAA	Terceira Island, Azores, Portugal	38.77°N	27.38°W	19	566	14	6	0.7–2.2	1.08	0.12	1.44
co2_azv_tower-insitu_20_allvalid-29magl	NIES	Azovo, Russia	54.70°N	73.03°E	110	12850	339	0	1.6–3.8	0.85	-0.31	2.95
co2_azv_tower-insitu_20_allvalid-50magl	NIES	Azovo, Russia	54.70°N	73.03°E	110	12566	339	0	1.5–3.7	0.89	-0.09	2.85
co2_bal_surface-flask_1_representative	NOAA	Baltic Sea, Poland	55.35°N	17.22°E	3	910	4	0	0.4–11.1	0.77	-1.02	5.21
co2_bao_surface-pfp_1_allvalid-300magl	NOAA	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1579	2279	4	0	2.1–15.6	0.33	-1.22	2.75
co2_bao_tower-insitu_1_allvalid-100magl	NOAA	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1579	10250	224	0	2.3–13.6	0.71	-3.46	7.06
co2_bao_tower-insitu_1_allvalid-22magl	NOAA	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1579	10359	234	0	2.4–17.1	0.71	-3.96	8.59
co2_bao_tower-insitu_1_allvalid-305magl	NOAA	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1579	67645	1169	0	1.9–15.1	0.64	-0.79	4.52
co2_bck_surface-insitu_6_allvalid	ECCC	Behchoko, Northwest Territories, Canada	62.80°N	115.92°W	160	12958	297	0	0.8–3.3	1.04	-0.08	1.88
co2_bcs_surface-flask_426_representative	SIO_CO2	Baja California Sur, Mexico	23.30°N	110.20°W	4	5	0	0	0.5–1.6	1.90	-1.51	1.35
co2_bgi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	612	4	0	0	3.1–4.2	0.07	-1.10	0.89
co2_bgi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	612	4	0	0	3.1–4.2	0.07	-1.10	0.89

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_bgi_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	1585	39	1	0	0.3–3.9	0.73	0.30	2.03
co2_bgi_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	2546	13	2	0	0.3–2.1	0.90	0.25	0.78
co2_bgi_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	3516	41	4	0	0.2–2.0	1.37	0.57	1.17
co2_bgi_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	4569	24	1	0	0.2–1.6	0.98	0.04	0.65
co2_bgi_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	5489	28	5	0	0.2–1.8	1.46	-0.21	0.60
co2_bgi_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	6469	33	0	0	0.4–2.1	1.01	0.11	0.85
co2_bgi_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	7463	27	2	0	0.2–1.0	2.19	-0.11	0.50
co2_bgi_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Bradgate, Iowa, United States	42.82°N	94.41°W	8050	3	0	0	0.5–0.5	3.86	0.22	0.60
co2_bhd_surface-flask_1_representative	NOAA	Baring Head Station, New Zealand	41.41°S	174.87°E	85	298	3	0	0.4–2.3	1.07	-0.17	0.97
co2_bkt_surface-flask_1_representative	NOAA	Bukit Kototabang, Indonesia	0.20°S	100.32°E	845	643	9	3	4.1–7.6	1.61	6.39	4.32
co2_bme_surface-flask_1_representative	NOAA	St. Davids Head, Bermuda, United Kingdom	32.37°N	64.65°W	12	209	8	0	0.7–3.0	1.10	0.53	1.67
co2_bmw_surface-flask_1_representative	NOAA	Tudor Hill, Bermuda, United Kingdom	32.26°N	64.88°W	30	861	28	5	0.7–1.9	1.36	0.77	1.43

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_bne_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	635	58	0	0	0.9–5.3	0.56	0.12	3.10
co2_bne_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	635	58	0	0	0.9–5.3	0.56	0.12	3.10
co2_bne_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	1379	149	5	0	0.4–5.2	0.78	0.14	2.26
co2_bne_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	2298	96	6	0	0.4–2.8	1.17	0.06	1.64
co2_bne_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	3399	131	6	0	0.2–2.2	1.33	0.03	1.07
co2_bne_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	4278	66	6	0	0.3–1.9	1.00	-0.00	0.88
co2_bne_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	5386	96	9	0	0.2–1.4	1.10	-0.14	0.76
co2_bne_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	6360	97	5	0	0.2–1.5	1.33	-0.02	0.79
co2_bne_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	7650	65	10	0	0.2–1.5	1.09	0.03	0.71
co2_bne_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Beaver Crossing, Nebraska, United States	40.80°N	97.18°W	8071	24	2	0	0.3–1.7	0.84	-0.18	0.76
co2_bra_surface-insitu_6_allvalid	ECCC	Bratt's Lake Saskatchewan, Canada	50.20°N	104.71°W	595	14354	198	0	1.5–4.1	0.76	-0.08	2.00

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_brw_surface-flask_1_representative	NOAA	Barrow Atmospheric Baseline Observatory, United States	71.32°N	156.61°W	11	2131	45	10	0.3–4.5	0.78	-0.17	0.89
co2_brw_surface-flask_426_representative	SIO_CO2	Barrow Atmospheric Baseline Observatory, United States	71.32°N	156.61°W	11	262	1	0	0.6–5.4	0.54	-0.24	0.99
co2_brw_surface-insitu_1_allvalid	NOAA	Barrow Atmospheric Baseline Observatory, United States	71.32°N	156.61°W	11	43135	1229	0	0.5–2.7	1.08	-0.07	0.68
co2_brz_aircraft-insitu_20_allvalid_0-1000masl	NIES	Berezorechka, Russia	56.15°N	84.33°E	636	31275	213	0	1.1–158.4	0.26	-0.25	2.90
co2_brz_aircraft-insitu_20_allvalid_0-1000masl	NIES	Berezorechka, Russia	56.15°N	84.33°E	636	31275	213	0	1.1–158.4	0.26	-0.25	2.90
co2_brz_aircraft-insitu_20_allvalid_1000-2000masl	NIES	Berezorechka, Russia	56.15°N	84.33°E	1501	42027	180	0	1.0–71.6	0.29	-0.35	2.31
co2_brz_aircraft-insitu_20_allvalid_2000-3000masl	NIES	Berezorechka, Russia	56.15°N	84.33°E	2408	18893	132	0	0.6–61.2	0.39	0.27	2.18
co2_brz_aircraft-insitu_20_allvalid_3000-4000masl	NIES	Berezorechka, Russia	56.15°N	84.33°E	3085	2653	19	0	0.4–43.5	0.38	0.41	2.35
co2_brz_tower-insitu_20_allvalid-20magl	NIES	Berezorechka, Russia	56.15°N	84.33°E	168	12409	319	0	1.7–10.2	0.92	0.21	3.30
co2_brz_tower-insitu_20_allvalid-40magl	NIES	Berezorechka, Russia	56.15°N	84.33°E	168	12045	297	0	1.7–9.4	0.92	0.04	3.31
co2_brz_tower-insitu_20_allvalid-5magl	NIES	Berezorechka, Russia	56.15°N	84.33°E	168	15580	399	0	1.7–7.4	0.89	0.01	3.31

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_brz_tower-insitu_20_allvalid-80magl	NIES	Berezorechka, Russia	56.15°N	84.33°E	168	11663	332	0	1.6–6.1	0.96	-0.11	3.03
co2_bsc_surface-flask_1_representative	NOAA	Black Sea, Constanta, Romania	44.18°N	28.66°E	0	404	1	0	1.9–18.1	0.78	-5.25	9.22
co2_car_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	1777	63	5	0	0.9–4.2	0.77	-0.50	1.96
co2_car_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	1777	63	5	0	0.9–4.2	0.77	-0.50	1.96
co2_car_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	2456	1139	27	12	0.6–3.6	0.93	0.02	1.44
co2_car_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	3455	1243	52	8	0.2–1.7	1.18	-0.00	0.87
co2_car_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	4497	1131	52	11	0.1–1.8	1.01	0.08	0.76
co2_car_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	5487	844	44	4	0.3–1.2	1.10	0.04	0.70
co2_car_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	6425	826	29	4	0.4–1.9	1.02	0.05	0.67
co2_car_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	7468	694	31	0	0.4–1.5	1.06	-0.03	0.66
co2_car_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	8222	158	8	0	0.2–1.7	1.26	-0.14	0.71

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_car_aircraft-pfp_1_allvalid_9000-10000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	9140	2	0	0	0.7–0.7	1.97	-0.46	1.03
co2_car_aircraft-pfp_1_allvalid_11000-12000masl	NOAA	Briggsdale, Colorado, United States	40.63°N	104.33°W	11869	4	0	0	0.6–0.8	0.98	-0.44	0.57
co2_cba_surface-flask_1_representative	NOAA	Cold Bay, Alaska, United States	55.21°N	162.72°W	21	1702	25	9	0.7–4.5	0.85	-0.70	1.59
co2_cba_surface-flask_4_representative	SIO	Cold Bay, Alaska, United States	55.21°N	162.72°W	21	491	10	3	0.2–7.6	0.96	-0.78	2.26
co2_cby_surface-insitu_6_allvalid	ECCC	Cambridge Bay, Nunavut Territory, Canada	69.13°N	105.06°W	35	9207	301	0	0.5–1.6	1.39	-0.04	0.93
co2_cdl_surface-insitu_6_allvalid	ECCC	Candle Lake, Saskatchewan, Canada	53.99°N	105.12°W	600	8608	205	0	1.2–3.0	0.99	0.13	1.73
co2_cfa_surface-flask_2_representative	CSIRO	Cape Ferguson, Queensland, Australia	19.28°S	147.06°E	2	631	5	0	0.1–2.3	0.60	-0.14	0.98
co2_cgo_surface-flask_1_representative	NOAA	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94	842	0	1	0.2–3.2	0.76	-0.02	0.42
co2_cgo_surface-flask_2_representative	CSIRO	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94	1284	6	0	0.3–2.4	0.49	-0.03	0.40
co2_cgo_surface-flask_4_representative	SIO	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94	491	10	4	0.3–2.2	0.72	-0.14	0.99
co2_chl_surface-insitu_6_allvalid	ECCC	Churchill, Manitoba, Canada	58.74°N	93.82°W	29	8345	123	0	0.7–2.5	1.32	-0.09	1.24
co2_chm_surface-insitu_6_allvalid	ECCC	Chibougamau, Quebec, Canada	49.69°N	74.34°W	393	3531	43	0	1.5–3.2	0.78	0.10	2.10
co2_chr_surface-flask_1_representative	NOAA	Christmas Island, Republic of Kiribati	1.70°N	157.15°W	0	590	0	0	0.3–1.5	1.12	0.08	0.71
co2_chr_surface-flask_426_representative	SIO_CO2	Christmas Island, Republic of Kiribati	1.70°N	157.15°W	0	96	0	0	0.5–1.8	0.92	0.09	1.05
co2_cib_surface-flask_1_representative	NOAA	Centro de Investigacion de la Baja Atmosfera (CIBA), Spain	41.81°N	4.93°W	845	574	6	2	2.7–8.7	0.74	0.86	3.41

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_cma_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	638	686	11	1	1.3–5.1	0.76	0.27	2.61
co2_cma_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	638	686	11	1	1.3–5.1	0.76	0.27	2.61
co2_cma_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	1531	407	12	1	0.4–4.0	0.87	-0.23	2.10
co2_cma_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	2301	508	14	2	0.6–3.5	0.81	-0.12	1.69
co2_cma_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	3451	481	17	1	0.4–1.9	1.07	0.02	1.14
co2_cma_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	4195	226	15	0	0.4–1.8	1.28	0.26	1.15
co2_cma_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	5339	449	14	1	0.6–2.3	0.79	-0.05	1.00
co2_cma_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	6270	370	13	1	0.3–2.0	1.06	-0.14	0.96
co2_cma_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	7726	344	16	1	0.4–1.6	1.15	-0.10	0.90
co2_cma_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Offshore Cape May, New Jersey, United States	38.83°N	74.32°W	8047	53	3	0	0.6–1.5	1.15	0.24	1.09

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_con_aircraft-flask_42_allvalid_0-1000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	719	2	0	0	0.1–0.1	3.11	-0.18	0.17
co2_con_aircraft-flask_42_allvalid_0-1000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	719	2	0	0	0.1–0.1	3.11	-0.18	0.17
co2_con_aircraft-flask_42_allvalid_1000-2000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	1556	4	0	0	0.1–0.6	2.84	0.39	0.75
co2_con_aircraft-flask_42_allvalid_3000-4000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	3463	1	1	0	0.2–0.2	1.50	-0.54	0.83
co2_con_aircraft-flask_42_allvalid_5000-6000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	5499	0	2	0	0.3–0.3	–	-1.45	0.51
co2_con_aircraft-flask_42_allvalid_6000-7000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)		variable	6485	5	0	0	0.1–1.6	1.63	-0.09	1.06

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_con_aircraft-flask_42_allvalid_8000-9000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)	variable		8504	6	5	0	0.1–1.0	2.13	-0.83	0.53
co2_con_aircraft-flask_42_allvalid_9000-10000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)	variable		9627	281	31	1	0.2–1.9	1.79	0.16	0.83
co2_con_aircraft-flask_42_allvalid_10000-11000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)	variable		10671	1812	239	2	0.1–1.9	1.65	0.05	0.92
co2_con_aircraft-flask_42_allvalid_11000-12000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)	variable		11572	1121	112	0	0.1–1.8	1.48	-0.21	1.10
co2_con_aircraft-flask_42_allvalid_12000-13000masl	NIES	CONTRAIL (Comprehensive Observation Network for TRace gases by AIrLiner)	variable		12283	280	32	0	0.1–2.2	1.13	-0.27	1.07
co2_cps_surface-insitu_6_allvalid	ECCC	Chapais,Quebec, Canada	49.82°N	74.98°W	391	13721	159	0	1.0–3.3	0.76	0.29	2.15
co2_cpt_surface-flask_1_representative	NOAA	Cape Point, South Africa	34.35°S	18.49°E	230	295	4	1	0.3–1.3	0.54	0.16	0.43
co2_cpt_surface-insitu_36_marine	SAWS	Cape Point, South Africa	34.35°S	18.49°E	230	131536	5233	0	0.3–1.2	0.95	0.02	0.55

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_crv_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	315	1379	28	0	0.4–47.3	0.89	-2.22	5.41
co2_crv_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	315	1379	28	0	0.4–47.3	0.89	-2.22	5.41
co2_crv_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	1438	77	3	0	0.3–7.4	0.93	-0.29	1.14
co2_crv_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	2555	64	8	0	0.2–2.2	1.29	-0.28	1.13
co2_crv_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	3390	50	7	0	0.3–1.4	1.42	-0.10	1.04
co2_crv_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	4518	27	4	0	0.3–2.1	1.87	0.37	1.22

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_crv_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	5269	267	16	0	0.2–1.8	1.79	0.16	1.36
co2_crv_surface-pfp_1_allvalid-32magl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	611	1947	26	29	0.5–7.6	0.84	-0.15	2.50
co2_crv_tower-insitu_1_allvalid-17magl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	611	12913	237	0	1.0–4.2	1.04	-0.06	2.39
co2_crv_tower-insitu_1_allvalid-32magl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	611	13500	222	0	0.9–5.6	1.04	-0.15	2.22
co2_crv_tower-insitu_1_allvalid-5magl	NOAA	Carbon in Arctic Reservoirs Vulnerability Experiment (CARVE), United States	64.99°N	147.60°W	611	13173	227	0	1.1–4.2	1.04	0.02	2.54
co2_crz_surface-flask_1_representative	NOAA	Crozet Island, France	46.43°S	51.85°E	197	918	0	0	0.2–0.4	0.71	-0.02	0.27
co2_cya_surface-flask_2_representative	CSIRO	Casey, Antarctica, Australia	66.28°S	110.52°E	47	626	0	0	0.1–0.6	0.76	0.03	0.25
co2_dem_tower-insitu_20_allvalid-45magl	NIES	Demyanskoe, Russia	59.79°N	70.87°E	63	16413	407	0	1.3–3.5	0.93	0.00	2.31

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_dem_tower-insitu_20_allvalid-63magl	NIES	Demyanskoe, Russia	59.79°N	70.87°E	63	15433	385	0	1.3–3.5	0.92	-0.11	2.32
co2_dnd_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	745	205	1	0	0.3–5.5	0.62	0.21	2.06
co2_dnd_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	745	205	1	0	0.3–5.5	0.62	0.21	2.06
co2_dnd_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	1502	277	5	0	0.1–4.5	0.76	-0.01	1.78
co2_dnd_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	2466	306	21	0	0.1–2.6	1.02	-0.04	1.37
co2_dnd_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	3531	232	10	0	0.4–3.1	1.36	-0.09	1.16
co2_dnd_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	4485	178	2	0	0.3–3.3	0.99	-0.01	0.93
co2_dnd_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	5503	200	8	0	0.4–2.0	0.82	-0.03	0.76
co2_dnd_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	6478	180	10	0	0.3–2.4	1.19	0.06	0.90
co2_dnd_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	7477	187	8	0	0.2–1.5	0.94	0.01	0.83
co2_dnd_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Dahlen, North Dakota, United States	47.50°N	99.24°W	8053	3	1	0	0.5–1.7	0.04	-0.16	1.20

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_drp_shipboard-flask_1_representative	NOAA	Drake Passage	variable		Surface	258	7	0	0.0–0.9	0.72	-0.06	0.28
co2_dsi_surface-flask_1_representative	NOAA	Dongsha Island, Taiwan	20.70°N	116.73°E	3	461	58	1	1.7–4.9	1.92	3.69	3.82
co2_egb_surface-insitu_6_allvalid	ECCC	Egbert, Ontario, Canada	44.23°N	79.78°W	251	16542	168	0	2.1–4.5	0.64	0.44	3.24
co2_eic_surface-flask_1_representative	NOAA	Easter Island, Chile	27.16°S	109.43°W	47	553	3	0	0.5–1.7	1.11	0.39	0.94
co2_esp_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	565	908	14	7	0.3–4.3	0.86	-0.30	2.28
co2_esp_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	565	908	14	7	0.3–4.3	0.86	-0.30	2.28
co2_esp_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	1578	1048	60	7	0.3–2.1	1.17	-0.12	1.23
co2_esp_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	2553	983	66	9	0.2–2.4	1.20	-0.22	1.24
co2_esp_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	3550	866	75	6	0.2–2.1	1.26	-0.12	1.14
co2_esp_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	4519	748	39	6	0.3–2.0	1.25	-0.08	1.13
co2_esp_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	5370	515	27	4	0.5–1.6	1.18	-0.06	1.07
co2_esp_surface-flask_2_representative	CSIRO	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	7	19	0	0	0.2–6.3	0.66	-1.22	1.37
co2_esp_surface-insitu_6_allvalid	ECCC	Estevan Point, British Columbia, Canada	49.38°N	126.54°W	7	14171	378	0	1.2–3.1	0.90	-0.23	1.78

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_est_surface-insitu_6_allvalid	ECCC	Esther, Alberta, Canada	51.67°N	110.21°W	707	15450	187	0	1.6–3.9	0.72	0.09	2.27
co2_etl_aircraft-pfp_1_allvalid_0-1000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	878	271	5	8	0.7–2.6	0.80	-0.31	1.62
co2_etl_aircraft-pfp_1_allvalid_0-1000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	878	271	5	8	0.7–2.6	0.80	-0.31	1.62
co2_etl_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	1489	895	20	12	0.6–3.3	0.88	-0.19	1.49
co2_etl_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	2482	906	12	9	0.7–6.5	1.09	-0.25	1.21
co2_etl_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	3484	360	9	6	0.6–1.8	1.23	-0.17	1.18
co2_etl_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	4578	257	8	0	0.6–5.4	1.05	0.16	1.37
co2_etl_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	5642	236	9	0	0.5–1.5	1.19	0.18	1.18
co2_etl_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	6761	130	2	0	0.5–1.8	0.70	0.17	0.79
co2_etl_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	7151	89	3	0	0.6–2.0	1.30	0.71	1.37
co2_etl_surface-insitu_6_allvalid	ECCC	East Trout Lake, Saskatchewan, Canada	54.35°N	104.99°W	493	22015	384	0	0.9–3.4	0.85	-0.08	1.83
co2_fsd_surface-insitu_6_allvalid	ECCC	Fraserdale, Canada	49.88°N	81.57°W	210	26951	333	0	1.1–3.2	0.73	0.02	1.96

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_ftl_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Fortaleza, Brazil	3.52°S	38.28°W	1810	8	2	0	0.3–0.7	1.80	0.50	0.38
co2_ftl_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Fortaleza, Brazil	3.52°S	38.28°W	1810	8	2	0	0.3–0.7	1.80	0.50	0.38
co2_ftl_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Fortaleza, Brazil	3.52°S	38.28°W	2498	21	2	0	0.3–1.5	1.62	0.16	0.93
co2_ftl_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Fortaleza, Brazil	3.52°S	38.28°W	3479	29	8	0	0.5–1.9	0.89	0.82	1.62
co2_ftl_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Fortaleza, Brazil	3.52°S	38.28°W	4267	6	1	0	0.5–2.8	1.60	0.34	1.05
co2_ftw_shipboard-insitu_20_allvalid	NIES	Fujitrans World (M/S Fujitrans World of Kagoshima Shipping Co., Ltd.)	variable		Surface	13207	881	0	0.1–9.9	1.00	-0.17	1.78
co2_ftws_shipboard-insitu_20_allvalid	NIES	Fujitrans World - Southeast Asia Route (M/S Fujitrans World of Kagoshima Shipping Co., Ltd.)	variable		Surface	42481	1992	0	0.1–17.1	0.90	0.20	3.25
co2_fwi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	623	7	0	0	2.5–4.0	0.83	1.87	2.53
co2_fwi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	623	7	0	0	2.5–4.0	0.83	1.87	2.53
co2_fwi_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	1570	41	4	0	0.3–4.8	1.22	0.38	1.98

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_fwi_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	2532	16	2	0	0.5–1.9	0.88	0.11	1.48
co2_fwi_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	3522	44	7	0	0.1–3.7	1.85	0.08	1.28
co2_fwi_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	4570	28	3	0	0.1–1.0	0.93	0.01	0.81
co2_fwi_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	5522	30	8	0	0.1–0.9	0.99	0.26	0.76
co2_fwi_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	6500	27	4	0	0.1–2.5	1.11	0.07	1.16
co2_fwi_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Fairchild, Wisconsin, United States	44.66°N	90.96°W	7499	31	4	0	0.3–2.5	1.41	0.34	1.42
co2_gmi_surface-flask_1_representative	NOAA	Mariana Islands, Guam	13.39°N	144.66°E	0	1108	56	0	0.2–1.2	1.13	0.16	0.87
co2_gw_shipboard-insitu_20_allvalid	NIES	Golden Wattle (M/S Alligator Hope of Mitsui O.S.K. Lines, Ltd.)	variable		Surface	1912	416	0	0.0–11.9	1.40	-0.03	1.66
co2_haa_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	890	18	0	0	0.3–1.0	1.65	0.25	0.67
co2_haa_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	890	18	0	0	0.3–1.0	1.65	0.25	0.67
co2_haa_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	1595	182	22	0	0.2–1.0	1.78	-0.09	0.94

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_haa_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	2527	192	8	0	0.2–1.3	1.47	-0.13	0.78
co2_haa_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	3488	195	17	0	0.2–1.2	1.33	-0.10	0.69
co2_haa_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	4531	221	18	0	0.1–1.0	1.49	-0.13	0.76
co2_haa_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	5441	166	19	0	0.1–1.0	1.57	-0.07	0.78
co2_haa_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	6480	179	9	0	0.1–1.5	1.45	0.09	0.97
co2_haa_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	7470	97	4	0	0.3–1.8	1.09	0.04	1.08
co2_haa_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Molokai Island, Hawaii, United States	21.23°N	158.95°W	8041	56	2	0	0.2–0.9	1.42	-0.18	0.61
co2_hba_surface-flask_1_representative	NOAA	Halley Station, Antarctica, United Kingdom	75.55°S	25.63°W	30	768	0	0	0.1–0.4	1.26	0.02	0.17
co2_hdp_surface-insitu_3_nonlocal	NCAR	Hidden Peak (Snowbird), Utah, United States	40.56°N	111.65°W	3351	55214	1964	0	0.7–2.0	1.11	-0.31	1.22
co2_hfm_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	766	139	0	0	0.3–6.1	0.72	-0.14	2.54
co2_hfm_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	766	139	0	0	0.3–6.1	0.72	-0.14	2.54

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_hfm_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	1531	238	2	0	0.7–5.6	0.94	-0.02	2.36
co2_hfm_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	2454	178	8	0	0.6–8.8	0.82	-0.29	2.51
co2_hfm_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	3438	146	7	0	0.3–5.2	0.79	0.21	1.22
co2_hfm_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	4565	157	16	0	0.4–2.2	0.82	0.23	1.09
co2_hfm_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	5476	170	20	0	0.1–1.4	0.77	0.18	0.91
co2_hfm_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	6425	130	7	0	0.2–1.7	0.93	0.16	0.95
co2_hfm_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	7389	148	4	0	0.3–3.0	0.68	0.05	0.75
co2_hfm_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Harvard Forest, Massachusetts, United States	42.54°N	72.17°W	8031	2	0	0	1.4–1.4	0.33	-0.12	1.06
co2_hil_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	628	252	0	1	1.0–5.9	0.59	0.06	2.50
co2_hil_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	628	252	0	1	1.0–5.9	0.59	0.06	2.50
co2_hil_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	1538	613	4	5	0.7–6.5	0.75	0.13	2.38

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_hil_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	2543	361	7	3	0.1–3.9	0.88	-0.15	1.40
co2_hil_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	3495	585	17	3	0.3–2.9	1.00	-0.18	1.21
co2_hil_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	4531	409	11	4	0.4–2.5	0.82	-0.20	1.01
co2_hil_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	5506	488	15	1	0.1–2.1	0.85	-0.17	0.91
co2_hil_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	6526	428	11	2	0.2–1.9	0.89	-0.23	0.99
co2_hil_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	7494	458	15	2	0.4–1.9	0.94	-0.27	0.96
co2_hil_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Homer, Illinois, United States	40.07°N	87.91°W	8044	28	1	0	0.3–2.7	0.73	-0.72	1.26
co2_hpb_surface-flask_1_representative	NOAA	Hohenpeissenberg, Germany	47.80°N	11.02°E	985	785	19	3	3.5–11.1	1.29	4.17	7.06
co2_hun_surface-flask_1_representative	NOAA	Hegyhatsal, Hungary	46.96°N	16.65°E	248	1139	7	0	2.2–8.3	0.55	-0.63	5.27
co2_ice_surface-flask_1_representative	NOAA	Storhofdi, Vestmannaeyjar, Iceland	63.40°N	20.29°W	118	721	30	5	0.1–1.9	0.93	0.05	1.07
co2_igr_tower-insitu_20_allvalid-24magl	NIES	Igrim, Russia	63.19°N	64.41°E	9	10777	227	0	2.9–5.5	0.64	-1.71	3.93
co2_igr_tower-insitu_20_allvalid-47magl	NIES	Igrim, Russia	63.19°N	64.41°E	9	10614	253	0	3.2–7.8	0.49	-1.58	5.53

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_inu_surface-insitu_6_allvalid	ECCC	Inuvik,Northwest Territories, Canada	68.32°N	133.53°W	113	14904	237	0	0.7–3.5	0.98	-0.09	2.00
co2_inx_aircraft-pfp_1_allvalid_0-1000masl	NOAA	INFLUX (Indianapolis Flux Experiment), United States	39.58°N	86.42°W	652	164	5	0	0.4–9.6	0.45	-1.69	3.41
co2_inx_aircraft-pfp_1_allvalid_0-1000masl	NOAA	INFLUX (Indianapolis Flux Experiment), United States	39.58°N	86.42°W	652	164	5	0	0.4–9.6	0.45	-1.69	3.41
co2_inx_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	INFLUX (Indianapolis Flux Experiment), United States	39.58°N	86.42°W	1354	56	1	0	0.5–5.5	0.74	0.09	1.93
co2_inx_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	INFLUX (Indianapolis Flux Experiment), United States	39.58°N	86.42°W	2501	20	1	0	0.6–2.1	1.05	0.08	1.22
co2_inx_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	INFLUX (Indianapolis Flux Experiment), United States	39.58°N	86.42°W	3226	7	2	0	0.3–1.0	1.48	0.10	0.67
co2_jfj_surface-insitu_5_allvalid	EMPA	Jungfraujoch, Switzerland	46.55°N	7.99°E	3570	13114	642	0	1.2–2.0	0.95	0.04	1.71
co2_key_surface-flask_1_representative	NOAA	Key Biscayne, Florida, United States	25.67°N	80.16°W	1	801	12	2	0.9–3.9	0.81	0.47	1.65
co2_krs_tower-insitu_20_allvalid-35magl	NIES	Karasevoe, Russia	58.25°N	82.42°E	76	15604	418	0	1.3–4.3	0.89	0.05	2.85
co2_krs_tower-insitu_20_allvalid-67magl	NIES	Karasevoe, Russia	58.25°N	82.42°E	76	14877	388	0	1.3–4.3	0.90	-0.10	2.79
co2_kum_surface-flask_1_representative	NOAA	Cape Kumukahi, Hawaii, United States	19.56°N	154.89°W	8	2067	24	11	0.4–2.4	0.60	-0.36	0.96
co2_kum_surface-flask_4_representative	SIO	Cape Kumukahi, Hawaii, United States	19.56°N	154.89°W	8	638	19	6	0.3–3.7	0.70	-0.45	1.23
co2_kum_surface-flask_426_representative	SIO_CO2	Cape Kumukahi, Hawaii, United States	19.56°N	154.89°W	8	2	0	0	0.3–0.3	0.38	-0.29	0.08

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_kzd_surface-flask_1_representative	NOAA	Sary Taukum, Kazakhstan	44.08°N	76.87°E	595	409	5	0	0.7–6.6	0.85	-0.82	3.53
co2_kzm_surface-flask_1_representative	NOAA	Plateau Assy, Kazakhstan	43.25°N	77.88°E	2519	364	4	0	0.9–4.8	1.00	0.10	2.38
co2_lef_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	781	753	5	6	0.7–7.0	0.89	0.02	2.04
co2_lef_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	781	753	5	6	0.7–7.0	0.89	0.02	2.04
co2_lef_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	1524	1372	25	8	0.4–3.9	0.85	-0.04	1.94
co2_lef_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	2468	1033	35	7	0.4–3.7	1.01	-0.10	1.71
co2_lef_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	3500	1206	41	5	0.4–2.9	1.04	0.00	1.45
co2_lef_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	4017	2	0	0	1.0–1.0	0.97	0.40	0.60
co2_lef_tower-insitu_1_allvalid-11magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	10719	177	0	1.1–4.3	0.89	0.44	2.53
co2_lef_tower-insitu_1_allvalid-122magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	30622	376	0	1.3–4.3	0.77	0.33	2.58
co2_lef_tower-insitu_1_allvalid-244magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	64761	869	0	1.0–6.5	0.81	-0.07	2.78
co2_lef_tower-insitu_1_allvalid-30magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	29416	331	0	1.4–4.6	0.80	0.49	2.70

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_lef_tower-insitu_1_allvalid-396magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	184350	2434	0	1.3–5.1	0.70	0.06	2.77
co2_lef_tower-insitu_1_allvalid-76magl	NOAA	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	10620	178	0	1.1–3.6	0.88	0.21	2.19
co2_lew_surface-pfp_1_allvalid-95magl	NOAA	Lewisburg, Pennsylvania, United States	40.94°N	76.88°W	166	1013	10	0	2.7–6.7	0.57	-0.56	4.10
co2_ljo_surface-flask_426_representative	SIO_CO2	La Jolla, California, United States	32.87°N	117.26°W	10	21	2	0	0.4–3.9	1.62	1.16	1.91
co2_llb_surface-flask_1_representative	NOAA	Lac La Biche, Alberta, Canada	54.95°N	112.47°W	540	146	0	0	1.1–11.1	0.52	-1.32	3.48
co2_llb_surface-insitu_6_allvalid	ECCC	Lac La Biche, Alberta, Canada	54.95°N	112.47°W	540	15872	158	0	1.5–5.7	0.60	-0.56	2.77
co2_lmp_surface-flask_1_representative	NOAA	Lampedusa, Italy	35.52°N	12.63°E	45	734	24	1	1.0–2.7	1.16	0.70	1.85
co2_lut_surface-insitu_44_allvalid	RUG	Lutjewad, Netherlands	53.40°N	6.35°E	1	21740	367	0	3.3–7.2	0.62	-0.05	5.06
co2_maa_surface-flask_2_representative	CSIRO	Mawson Station, Antarctica, Australia	67.62°S	62.87°E	32	723	0	0	0.1–0.8	1.00	0.04	0.26
co2_man_aircraft-insitu_1_allvalid_0-1000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	653	1915	37	0	1.8–7.9	1.16	0.51	2.40
co2_man_aircraft-insitu_1_allvalid_0-1000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	653	1915	37	0	1.8–7.9	1.16	0.51	2.40
co2_man_aircraft-insitu_1_allvalid_1000-2000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	1514	2639	71	0	0.9–3.1	1.12	0.00	1.04
co2_man_aircraft-insitu_1_allvalid_2000-3000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	2513	2569	131	0	0.4–2.2	1.06	0.05	0.87

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_man_aircraft-insitu_1_allvalid_3000-4000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	3513	2484	208	0	0.4–2.1	1.23	0.08	0.86
co2_man_aircraft-insitu_1_allvalid_4000-5000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	4498	2318	236	0	0.1–1.7	1.46	0.21	0.76
co2_man_aircraft-insitu_1_allvalid_5000-6000masl	NOAA	Manaus, Brazil	2.60°S	60.21°W	5181	584	65	0	0.4–1.6	1.75	0.26	0.90
co2_mbo_surface-pfp_1_allvalid-11magl	NOAA	Mt. Bachelor Observatory, United States	43.98°N	121.69°W	2731	2195	58	4	0.9–2.0	1.08	-0.28	1.53
co2_mex_surface-flask_1_representative	NOAA	High Altitude Global Climate Observation Center, Mexico	18.98°N	97.31°W	4464	421	15	0	0.9–2.9	0.79	0.61	1.44
co2_mhd_surface-flask_1_representative	NOAA	Mace Head, County Galway, Ireland	53.33°N	9.90°W	5	929	26	6	0.6–3.7	0.97	0.07	1.40
co2_mid_surface-flask_1_representative	NOAA	Sand Island, Midway, United States	28.22°N	177.37°W	5	1060	30	5	0.3–1.3	1.13	0.15	0.95
co2_mkn_surface-flask_1_representative	NOAA	Mt. Kenya, Kenya	0.06°S	37.30°E	3644	127	0	0	0.8–3.8	1.31	2.19	1.89
co2_mlo_surface-flask_1_representative	NOAA	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	2246	122	6	0.4–1.1	1.34	-0.18	0.63
co2_mlo_surface-flask_2_representative	CSIRO	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	1119	11	0	0.4–2.7	0.70	-0.03	0.67
co2_mlo_surface-flask_4_representative	SIO	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	871	38	4	0.2–1.3	1.14	-0.12	0.66
co2_mlo_surface-flask_426_representative	SIO_CO2	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	222	3	0	0.3–1.0	0.78	-0.33	0.59
co2_mlo_surface-insitu_1_allvalid	NOAA	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	45003	0	0	0.3–0.6	1.99	-0.13	0.55
co2_mqa_surface-flask_2_representative	CSIRO	Macquarie Island, Australia	54.48°S	158.97°E	6	774	0	0	0.2–1.0	0.62	0.22	0.34

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_mvy_surface-insitu_1_allvalid	NOAA	Marthas Vineyard, Massachusetts, United States	41.33°N	70.57°W	0	63073	0	0	2.1–8.9	1.05	-0.19	4.13
co2_mwo_surface-pfp_1_allvalid-46magl	NOAA	Mt. Wilson Observatory, United States	34.22°N	118.06°W	1729	3591	94	8	1.2–2.9	0.84	-0.79	1.55
co2_nat_surface-flask_1_representative	NOAA	Farol De Mae Luiza Lighthouse, Brazil	5.80°S	35.19°W	50	305	7	0	0.7–1.6	0.95	-0.06	1.02
co2_nat_surface-flask_26_marine	IPEN	Farol De Mae Luiza Lighthouse, Brazil	5.80°S	35.19°W	50	189	4	0	0.5–1.9	1.05	-0.03	1.09
co2_nha_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	616	923	7	0	0.5–5.1	0.85	0.06	2.35
co2_nha_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	616	923	7	0	0.5–5.1	0.85	0.06	2.35
co2_nha_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	1500	688	21	0	0.6–3.7	0.93	-0.09	2.10
co2_nha_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	2399	696	25	0	0.5–3.5	0.91	-0.21	1.59
co2_nha_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	3480	572	23	0	0.5–4.1	1.01	0.08	1.28
co2_nha_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	4408	321	19	0	0.2–2.0	0.97	0.13	1.20

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_nha_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	5368	453	16	0	0.4–2.1	0.84	0.01	1.06
co2_nha_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	6279	327	10	0	0.4–2.3	0.98	-0.04	1.15
co2_nha_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	7662	360	8	0	0.3–1.7	0.95	0.03	1.05
co2_nha_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Offshore Portsmouth, New Hampshire (Isles of Shoals), United States	42.95°N	70.63°W	8038	5	1	0	0.1–1.7	0.88	-0.16	0.89
co2_nmb_surface-flask_1_representative	NOAA	Gobabeb, Namibia	23.58°S	15.03°E	456	647	24	0	0.4–1.8	1.26	-0.03	1.43
co2_noy_tower-insitu_20_allvalid-21magl	NIES	Noyabrsk, Russia	63.43°N	75.78°E	108	11400	323	0	1.0–3.0	1.01	0.05	2.11
co2_noy_tower-insitu_20_allvalid-43magl	NIES	Noyabrsk, Russia	63.43°N	75.78°E	108	11752	347	0	1.0–3.0	1.03	-0.07	2.07
co2_nwr_surface-flask_1_representative	NOAA	Niwot Ridge, Colorado, United States	40.05°N	105.59°W	3523	1096	22	6	0.4–6.7	0.79	0.36	1.40
co2_nwr_surface-insitu_3_nonlocal	NCAR	Niwot Ridge, Colorado, United States	40.05°N	105.59°W	3523	63639	1511	0	0.7–3.9	0.81	0.05	1.34
co2_nwr_surface-pfp_1_allvalid-3magl	NOAA	Niwot Ridge, Colorado, United States	40.05°N	105.59°W	3523	4324	70	41	0.4–8.0	0.64	0.40	1.51
co2_obn_surface-flask_1_representative	NOAA	Obninsk, Russia	55.11°N	36.60°E	183	170	1	0	1.8–10.2	0.82	-0.06	5.51
co2_ofr_surface-insitu_68_allhours	OSU	Fir, Oregon, United States	44.65°N	123.55°W	263	55196	28	0	3.7–33.6	0.20	-1.65	6.33

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_omp_surface-insitu_68_allhours	OSU	Marys Peak, Oregon, United States	44.50°N	123.55°W	1249	124843	2728	0	1.1–18.9	0.68	0.92	2.80
co2_omt_surface-insitu_68_allhours	OSU	Metolius, Oregon, United States	44.45°N	121.56°W	1255	74729	399	0	2.0–22.9	0.44	1.24	3.55
co2_ong_surface-insitu_68_allhours	OSU	Burns, Oregon, United States	43.47°N	119.69°W	1398	80744	601	0	1.1–19.0	0.45	-0.35	2.73
co2_owa_surface-insitu_68_allhours	OSU	Walton, Oregon, United States	44.07°N	123.63°W	715	37272	571	0	1.8–6.8	0.85	-1.09	3.19
co2_oxk_surface-flask_1_representative	NOAA	Ochsenkopf, Germany	50.03°N	11.81°E	1022	645	18	2	0.5–4.9	0.85	-0.52	3.89
co2_oyq_surface-insitu_68_allhours	OSU	Yaquina Head, Oregon, United States	44.67°N	124.07°W	116	33355	79	0	1.1–35.1	0.26	-2.05	3.99
co2_pal_surface-flask_1_representative	NOAA	Pallas-Sammaltunturi, GAW Station, Finland	67.97°N	24.12°E	565	920	12	2	1.0–7.2	0.73	-0.63	2.39
co2_pal_surface-insitu_30_marine	FMI	Pallas-Sammaltunturi, GAW Station, Finland	67.97°N	24.12°E	565	21920	1126	0	0.4–1.9	1.13	-0.02	0.81
co2_pfa_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	553	778	18	8	0.3–3.5	1.05	-0.17	2.19
co2_pfa_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	553	778	18	8	0.3–3.5	1.05	-0.17	2.19
co2_pfa_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	1522	717	28	6	0.5–6.1	1.01	-0.43	1.33
co2_pfa_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	2528	767	40	5	0.6–3.4	1.21	-0.46	1.31
co2_pfa_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	3462	726	36	3	0.3–2.5	1.35	-0.26	1.33
co2_pfa_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	4494	637	33	3	0.2–1.8	1.29	-0.14	1.06

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_pfa_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	5436	599	28	3	0.3–2.1	1.26	-0.10	1.16
co2_pfa_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	6444	571	32	3	0.4–2.3	1.21	0.01	1.35
co2_pfa_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Poker Flat, Alaska, United States	64.90°N	148.76°W	7214	263	7	1	0.3–2.5	1.27	0.12	1.37
co2_poc_shipboard-flask_1_representative	NOAA	Pacific Ocean	variable		Surface	2027	120	0	0.1–1.9	1.09	-0.04	0.65
co2_prs_surface-insitu_21_allvalid	RSE	Plateau Rosa Station, Italy	45.93°N	7.70°E	3480	20597	1084	260	1.1–1.7	1.05	0.07	1.67
co2_psa_surface-flask_1_representative	NOAA	Palmer Station, Antarctica, United States	64.77°S	64.05°W	10	1149	0	0	0.1–0.5	0.72	-0.07	0.22
co2_psa_surface-flask_4_representative	SIO	Palmer Station, Antarctica, United States	64.77°S	64.05°W	10	534	8	0	0.1–2.1	0.61	-0.10	0.42
co2_pta_surface-flask_1_representative	NOAA	Point Arena, California, United States	38.95°N	123.74°W	17	372	2	0	1.8–9.2	0.92	-2.64	3.72
co2_px_shipboard-insitu_20_allvalid	NIES	Pyxis (M/S Pyxis of Toyofuji Shipping Co., Ltd.)	variable		Surface	62089	3979	0	0.1–11.6	1.13	0.12	1.75
co2_rpb_surface-flask_1_representative	NOAA	Ragged Point, Barbados	13.16°N	59.43°W	15	1075	23	0	0.2–1.1	1.03	0.21	0.74
co2_rta_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	685	107	7	0	0.2–0.7	1.18	0.31	0.46
co2_rta_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	685	107	7	0	0.2–0.7	1.18	0.31	0.46

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_rta_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	1654	342	41	7	0.2–1.0	1.90	0.43	0.56
co2_rta_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	2600	312	38	5	0.3–0.8	1.55	0.29	0.54
co2_rta_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	3481	479	62	7	0.2–0.9	1.62	0.21	0.68
co2_rta_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	4528	315	59	4	0.2–0.8	1.93	0.27	0.69
co2_rta_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	5460	401	71	7	0.2–1.0	1.91	0.22	0.74
co2_rta_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Rarotonga, Cook Islands	21.25°S	159.83°W	6289	302	32	4	0.3–1.5	1.57	0.37	0.71
co2_san_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Santarem, Brazil	2.85°S	54.95°W	1713	11	0	0	0.4–2.9	0.71	-1.24	2.11
co2_san_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Santarem, Brazil	2.85°S	54.95°W	1713	11	0	0	0.4–2.9	0.71	-1.24	2.11
co2_san_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Santarem, Brazil	2.85°S	54.95°W	2527	42	4	0	0.4–2.2	1.61	0.29	2.08
co2_san_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Santarem, Brazil	2.85°S	54.95°W	3436	65	8	0	0.2–2.1	1.90	0.52	1.61
co2_san_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Santarem, Brazil	2.85°S	54.95°W	4600	6	0	0	1.4–1.6	0.39	-0.62	0.89

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_sca_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	655	358	0	6	0.5–3.7	0.91	0.36	2.20
co2_sca_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	655	358	0	6	0.5–3.7	0.91	0.36	2.20
co2_sca_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	1531	273	7	4	0.5–2.9	0.82	0.48	1.74
co2_sca_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	2475	621	22	8	0.3–2.1	0.95	0.00	1.24
co2_sca_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	3543	435	12	7	0.2–1.2	0.94	0.01	0.83
co2_sca_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	4462	542	16	4	0.1–1.6	0.97	-0.10	0.80
co2_sca_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	5479	347	17	4	0.3–1.2	1.12	-0.08	0.81
co2_sca_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	6422	411	19	10	0.2–1.4	1.04	-0.12	0.84
co2_sca_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	7460	477	27	2	0.2–1.6	1.21	-0.17	0.83
co2_sca_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	8101	99	1	0	0.1–1.3	0.98	0.10	0.71
co2_sca_aircraft-pfp_1_allvalid_9000-10000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	9362	15	0	0	0.3–4.9	2.11	-0.79	2.02

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_sca_aircraft-pfp_1_allvalid_10000-11000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	10432	13	4	0	0.2–5.2	1.66	-0.97	1.70
co2_sca_aircraft-pfp_1_allvalid_11000-12000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	11160	5	0	0	1.0–5.7	0.93	-1.45	3.43
co2_sca_aircraft-pfp_1_allvalid_12000-13000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	12636	13	1	0	0.3–1.3	0.75	-0.09	0.51
co2_sca_aircraft-pfp_1_allvalid_13000-14000masl	NOAA	Offshore Charleston, South Carolina, United States	32.77°N	79.55°W	13145	2	0	0	0.6–0.6	0.70	0.05	0.68
co2_sct_surface-pfp_1_allvalid-305magl	NOAA	Beech Island, South Carolina, United States	33.41°N	81.83°W	115	1740	3	0	2.5–13.1	0.42	0.86	2.69
co2_sct_tower-insitu_1_allvalid-305magl	NOAA	Beech Island, South Carolina, United States	33.41°N	81.83°W	115	121660	1235	0	2.7–12.2	0.71	0.55	4.43
co2_sct_tower-insitu_1_allvalid-31magl	NOAA	Beech Island, South Carolina, United States	33.41°N	81.83°W	115	19261	248	0	3.2–4.7	0.67	0.72	3.80
co2_sct_tower-insitu_1_allvalid-61magl	NOAA	Beech Island, South Carolina, United States	33.41°N	81.83°W	115	20042	266	0	3.1–4.6	0.66	0.53	3.67
co2_sey_surface-flask_1_representative	NOAA	Mahe Island, Seychelles	4.68°S	55.53°E	2	1018	0	2	0.3–1.1	2.19	0.34	1.19
co2_sgp_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	677	1144	12	11	0.8–6.7	0.77	0.94	2.31
co2_sgp_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	677	1144	12	11	0.8–6.7	0.77	0.94	2.31

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_sgp_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	1567	1409	29	11	0.6–3.5	0.78	0.40	1.79
co2_sgp_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	2470	1399	33	12	0.3–2.2	0.90	-0.07	1.21
co2_sgp_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	3478	1038	33	6	0.1–1.8	0.91	-0.10	0.94
co2_sgp_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	4606	707	11	6	0.4–1.9	0.97	-0.13	0.77
co2_sgp_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	5395	223	8	0	0.4–1.1	1.20	-0.17	0.72
co2_sgp_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	6464	6	2	0	0.5–0.5	2.23	0.58	0.80
co2_sgp_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	8062	2	0	0	1.3–1.3	0.05	-0.18	0.23
co2_sgp_aircraft-pfp_1_allvalid_9000-10000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	9686	7	1	0	0.7–0.7	0.87	-0.15	0.84
co2_sgp_aircraft-pfp_1_allvalid_11000-12000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	11321	2	0	0	0.5–0.5	0.08	0.05	0.40
co2_sgp_aircraft-pfp_1_allvalid_12000-13000masl	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	12858	4	0	0	0.5–0.5	0.13	-0.30	0.24
co2_sgp_surface-flask_1_representative	NOAA	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	314	1121	10	6	1.9–7.8	0.61	0.55	3.31

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_sgp_surface-insitu_64_allvalid-60magl	LBNL-ARM	Southern Great Plains, Oklahoma, United States	36.61°N	97.49°W	314	21422	180	0	2.6–13.8	0.60	1.17	3.91
co2_shm_surface-flask_1_representative	NOAA	Shemya Island, Alaska, United States	52.71°N	174.13°E	23	815	24	0	0.8–4.7	0.83	-0.37	1.75
co2_sis_surface-flask_2_representative	CSIRO	Shetland Islands, Scotland	60.09°N	1.25°W	30	76	11	0	0.2–2.5	1.39	0.93	1.21
co2_smo_surface-flask_1_representative	NOAA	Tutuila, American Samoa	14.25°S	170.56°W	42	1908	78	14	0.2–1.4	1.11	-0.07	0.43
co2_smo_surface-flask_4_representative	SIO	Tutuila, American Samoa	14.25°S	170.56°W	42	717	10	7	0.1–7.7	0.60	-0.00	1.01
co2_smo_surface-flask_426_representative	SIO_CO2	Tutuila, American Samoa	14.25°S	170.56°W	42	55	0	0	0.3–2.6	0.46	-0.29	0.49
co2_smo_surface-insitu_1_allvalid	NOAA	Tutuila, American Samoa	14.25°S	170.56°W	42	42686	0	0	0.1–0.8	2.81	-0.02	0.31
co2_spl_surface-insitu_3_nonlocal	NCAR	Storm Peak Laboratory (Desert Research Institute), United States	40.45°N	106.73°W	3210	62268	1352	0	0.9–2.5	0.96	-0.51	1.60
co2_spo_surface-flask_1_representative	NOAA	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	1705	6	0	0.0–0.8	0.55	0.00	0.14
co2_spo_surface-flask_4_representative	SIO	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	566	6	1	0.1–2.0	0.55	-0.01	0.27
co2_spo_surface-flask_426_representative	SIO_CO2	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	399	2	0	0.1–0.7	0.93	-0.04	0.23
co2_spo_surface-insitu_1_allvalid	NOAA	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	69076	0	0	0.1–0.5	1.77	0.00	0.11
co2_stm_surface-flask_1_representative	NOAA	Ocean Station M, Norway	66.00°N	2.00°E	0	758	47	0	0.4–3.0	1.00	0.21	1.27
co2_sum_surface-flask_1_representative	NOAA	Summit, Greenland	72.60°N	38.42°W	3210	1031	31	0	0.2–1.2	1.06	-0.10	0.80
co2_svv_tower-insitu_20_allvalid-27magl	NIES	Savvushka, Russia	51.33°N	82.13°E	495	7591	212	0	1.3–3.5	0.99	-0.17	2.45

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_svv_tower-insitu_20_allvalid-52magl	NIES	Savvushka, Russia	51.33°N	82.13°E	495	7198	188	0	1.3–3.3	1.05	-0.16	2.45
co2_syo_surface-flask_1_representative	NOAA	Syowa Station, Antarctica, Japan	69.01°S	39.59°E	14	528	0	0	0.0–0.4	1.13	-0.09	0.18
co2_tap_surface-flask_1_representative	NOAA	Tae-ahn Peninsula, Republic of Korea	36.74°N	126.13°E	16	1404	19	2	0.8–10.2	0.77	0.87	5.94
co2_tf1_shipboard-insitu_20_allvalid	NIES	Trans Future 1 (M/S Trans Future 1 of the Toyofuji Shipping Co., Ltd)	variable		Surface	13777	913	0	0.4–12.4	1.05	0.33	3.71
co2_tf5_shipboard-insitu_20_allvalid	NIES	Trans Future 5 (M/S Trans Future 5 of Toyofuji Shipping Co., Ltd.)	variable		Surface	97061	5116	0	0.1–18.8	0.75	-0.08	2.84
co2_tgc_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	698	222	2	3	0.9–2.9	1.04	0.52	1.70
co2_tgc_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	698	222	2	3	0.9–2.9	1.04	0.52	1.70
co2_tgc_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	1557	242	5	2	0.5–2.2	0.96	0.26	1.25
co2_tgc_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	2530	565	15	4	0.4–1.4	0.93	0.03	0.95
co2_tgc_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	3493	316	11	2	0.2–1.2	0.85	-0.04	0.68
co2_tgc_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	4467	556	22	3	0.3–1.2	1.08	-0.17	0.70

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_tgc_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	5570	274	18	2	0.2–1.1	1.27	-0.22	0.64
co2_tgc_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	6418	296	19	3	0.1–1.1	1.55	-0.21	0.74
co2_tgc_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	7414	383	29	3	0.1–1.1	1.44	-0.21	0.71
co2_tgc_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Offshore Corpus Christi, Texas, United States	27.73°N	96.86°W	8083	110	14	0	0.3–1.0	1.53	-0.09	0.83
co2_thd_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	628	499	3	6	1.3–5.1	0.62	-0.70	2.97
co2_thd_aircraft-pfp_1_allvalid_0-1000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	628	499	3	6	1.3–5.1	0.62	-0.70	2.97
co2_thd_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	1526	284	7	3	0.5–2.2	1.33	-0.07	1.39
co2_thd_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	2468	509	28	6	0.2–1.8	1.07	-0.12	1.14
co2_thd_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	3517	349	12	4	0.4–1.5	1.14	-0.09	0.95
co2_thd_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	4446	412	11	5	0.2–2.1	1.15	-0.10	1.03
co2_thd_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	5481	254	13	3	0.2–2.0	0.90	-0.15	0.91

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_thd_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	6445	325	17	4	0.4–2.4	1.10	-0.13	0.99
co2_thd_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	7479	325	20	5	0.2–1.3	1.21	-0.16	0.95
co2_thd_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	8045	27	0	0	1.0–1.3	0.87	-0.13	0.87
co2_thd_surface-flask_1_representative	NOAA	Trinidad Head, California, United States	41.05°N	124.15°W	107	629	5	0	1.4–7.6	0.78	-2.09	3.76
co2_tik_surface-flask_1_representative	NOAA	Hydrometeorological Observatory of Tiksi, Russia	71.60°N	128.89°E	19	284	2	0	0.9–7.5	0.97	-0.54	3.35
co2_ulb_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	1648	159	2	0	0.3–4.0	1.05	0.05	1.46
co2_ulb_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	1648	159	2	0	0.3–4.0	1.05	0.05	1.46
co2_ulb_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	2471	141	7	0	0.3–2.4	1.36	-0.13	1.28
co2_ulb_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	3478	148	7	0	0.2–2.2	1.06	-0.36	1.64
co2_ulb_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	4209	50	6	0	0.1–1.9	1.02	0.12	1.36
co2_ulb_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	Ulaanbaatar, Mongolia	47.40°N	106.00°E	5718	2	0	0	0.5–0.5	0.70	0.63	0.22
co2_ush_surface-flask_1_representative	NOAA	Ushuaia, Argentina	54.85°S	68.31°W	12	508	7	0	0.2–0.8	0.63	-0.21	0.45

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_uta_surface-flask_1_representative	NOAA	Wendover, Utah, United States	39.90°N	113.72°W	1327	1098	12	4	0.9–4.8	0.89	0.47	1.92
co2_uum_surface-flask_1_representative	NOAA	Ulaan Uul, Mongolia	44.45°N	111.10°E	1007	904	23	0	1.7–4.1	0.95	-0.71	3.81
co2_vgn_tower-insitu_20_allvalid-42magl	NIES	Vaganovo, Russia	54.50°N	62.32°E	192	12356	419	0	1.6–3.4	1.05	0.29	2.61
co2_vgn_tower-insitu_20_allvalid-85magl	NIES	Vaganovo, Russia	54.50°N	62.32°E	192	12223	431	0	1.6–3.6	1.02	0.25	2.88
co2_wbi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	652	222	4	3	1.2–11.5	0.49	0.17	2.82
co2_wbi_aircraft-pfp_1_allvalid_0-1000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	652	222	4	3	1.2–11.5	0.49	0.17	2.82
co2_wbi_aircraft-pfp_1_allvalid_1000-2000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	1529	550	11	8	0.4–5.9	0.69	0.22	2.21
co2_wbi_aircraft-pfp_1_allvalid_2000-3000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	2547	330	9	4	0.5–6.4	1.02	-0.15	1.40
co2_wbi_aircraft-pfp_1_allvalid_3000-4000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	3497	537	21	4	0.3–2.6	1.11	-0.02	1.25
co2_wbi_aircraft-pfp_1_allvalid_4000-5000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	4540	390	16	6	0.5–2.3	0.91	-0.11	0.97
co2_wbi_aircraft-pfp_1_allvalid_5000-6000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	5510	460	10	1	0.4–1.8	0.81	-0.12	0.89
co2_wbi_aircraft-pfp_1_allvalid_6000-7000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	6515	402	12	1	0.3–2.1	0.69	-0.11	0.77

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_wbi_aircraft-pfp_1_allvalid_7000-8000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	7501	435	14	1	0.5–1.5	0.73	-0.14	0.90
co2_wbi_aircraft-pfp_1_allvalid_8000-9000masl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	8052	42	0	0	0.6–1.2	1.27	0.01	0.74
co2_wbi_surface-pfp_1_allvalid-379magl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	242	2990	13	27	2.3–30.6	0.27	0.03	2.85
co2_wbi_tower-insitu_1_allvalid-31magl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	242	21163	216	0	2.9–7.5	0.43	0.49	4.52
co2_wbi_tower-insitu_1_allvalid-379magl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	242	117002	1300	0	2.4–11.5	0.49	0.21	4.21
co2_wbi_tower-insitu_1_allvalid-99magl	NOAA	West Branch, Iowa, United States	41.72°N	91.35°W	242	21923	214	0	2.6–7.2	0.42	0.32	4.33
co2_wgc_tower-insitu_1_allvalid-30magl	NOAA	Walnut Grove, California, United States	38.26°N	121.49°W	2	21213	390	0	2.7–15.3	0.80	-2.39	7.96
co2_wgc_tower-insitu_1_allvalid-484magl	NOAA	Walnut Grove, California, United States	38.26°N	121.49°W	2	112063	2800	0	2.0–7.5	0.69	-0.58	5.00
co2_wgc_tower-insitu_1_allvalid-89magl	NOAA	Walnut Grove, California, United States	38.26°N	121.49°W	2	21334	505	0	2.5–13.4	0.89	-2.62	7.57
co2_wis_surface-flask_1_representative	NOAA	Weizmann Institute of Science at the Arava Institute, Ketura, Israel	29.96°N	35.06°E	151	988	25	0	1.3–3.5	0.99	0.14	2.25
co2_wkt_tower-insitu_1_allvalid-122magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	25879	472	0	2.2–3.9	0.57	-0.28	2.87

(table continued on next page)

Dataset	Lab.	Location	Latitude	Longitude	Elev. (m)	Used	Rej.	Unsampled	R (ppm)	χ^2	Bias (ppm)	SE (ppm)
co2_wkt_tower-insitu_1_allvalid-244magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	15271	215	0	2.2–5.9	0.70	-0.60	3.34
co2_wkt_tower-insitu_1_allvalid-30magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	27152	450	0	2.4–4.1	0.57	-0.07	2.98
co2_wkt_tower-insitu_1_allvalid-457magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	131513	2248	0	2.1–5.4	0.68	0.20	2.84
co2_wkt_tower-insitu_1_allvalid-62magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	2532	44	0	1.8–4.2	0.58	-0.52	2.61
co2_wkt_tower-insitu_1_allvalid-9magl	NOAA	Moody, Texas, United States	31.31°N	97.33°W	251	2916	45	0	2.2–5.9	0.56	-0.51	3.74
co2_wlg_surface-flask_1_representative	NOAA	Mt. Waliguan, Peoples Republic of China	36.29°N	100.90°E	3810	1007	59	2	1.0–3.2	0.78	-0.19	3.04
co2_wpc_shipboard-flask_1_representative	NOAA	Western Pacific Cruise	variable		Surface	150	30	0	0.0–1.8	1.68	-0.07	0.84
co2_wsa_surface-insitu_6_allvalid	ECCC	Sable Island, Nova Scotia, Canada	43.93°N	60.01°W	5	19189	376	0	1.3–3.3	0.69	-0.15	1.99
co2_yak_tower-insitu_20_allvalid-11magl	NIES	Yakutsk, Russia	62.09°N	129.36°E	264	4903	110	0	0.9–5.6	0.81	-0.69	4.09
co2_yak_tower-insitu_20_allvalid-77magl	NIES	Yakutsk, Russia	62.09°N	129.36°E	264	5140	94	0	1.0–5.1	0.95	-0.03	3.72
co2_zep_surface-flask_1_representative	NOAA	Ny-Alesund, Svalbard, Norway and Sweden	78.91°N	11.89°E	474	1171	43	0	0.4–1.4	1.06	0.02	0.79
co2_zep_surface-insitu_56_allvalid	NILU	Ny-Alesund, Svalbard, Norway and Sweden	78.91°N	11.89°E	474	22846	798	243	0.4–1.2	1.08	0.15	0.83

Summary of Observational Sites Used in CarbonTracker. The site location is specified by latitude, longitude and elevation in meters above sea level. The number of observations actually assimilated for each dataset is listed in the column "Used", and the number rejected due to inability to fit the observations is listed in the column "Rej.". Model-data-mismatch (R) is a value assigned to a given site that is meant to quantify our expected ability to simulate observations there. In this table we report the range of R values assigned to dataset observations by our "adaptive" model-data mismatch scheme (Section 5.2). These values are principally determined from the limitations of the atmospheric transport model. It is part of the standard deviation used to interpret the difference between a simulation first guess (Hx) of an observation and the actual measured value (z). The other component, HPH^{T}, is a measure of the ability of the ensemble Kalman filter to improve its simulated value for this observation by adjusting fluxes. These elements together form the innovation χ statistic for the site: $\chi = (z - Hx)/\sqrt{(HPH^{\mathrm{T}} + R^2)}$. The innovation χ^2 reported above is the mean of all squared normalized values for a given site. An average χ^2 below 1.0 indicates that the $HPH^{\mathrm{T}} + R^2$ values are too large. Conversely, values above 1.0 mean that this standard deviation is underestimated. The bias and SE columns are statistics of the posterior residuals (final modeled values minus measured values). The bias is the mean of these residuals; the SE is the standard error of those residuals.

Appendix C

Ecoregions in CarbonTracker

C.1 What are ecoregions?

Ecoregions are the actual scale on which CarbonTracker performs its optimization over land. Ecoregions are meant to represent large expanses of land within a given continent having similar ecosystem types, and are used to divide continent-scale regions into smaller domains for analysis. The ecosystem types use in CarbonTracker are derived from the Olson et al. (1992) vegetation classification (Table C.1, Figure C.1).

We define an ecoregion as an ecosystem type within a given Transcom land region. There are 19 ecosystem types we extract from the Olson et al. (1992) system, and 11 Transcom land regions (Figure C.2), so there are 11 × 19 = 209 possible ecoregions. However, not all ecosystem types are present in all Transcom regions, and the actual number of land ecoregions ends up being 126.

Note on "Semitundra": this is a potentially misleading shorthand abbreviation for a collection of ecosystems comprising semi-desert, shrubs, steppe, and polar+alpine tundra. The "Semitundra" zones appearing in northern Africa where one expects to find the Sahara desert are not, of course, tundra environments. They are instead semi-desert zones.

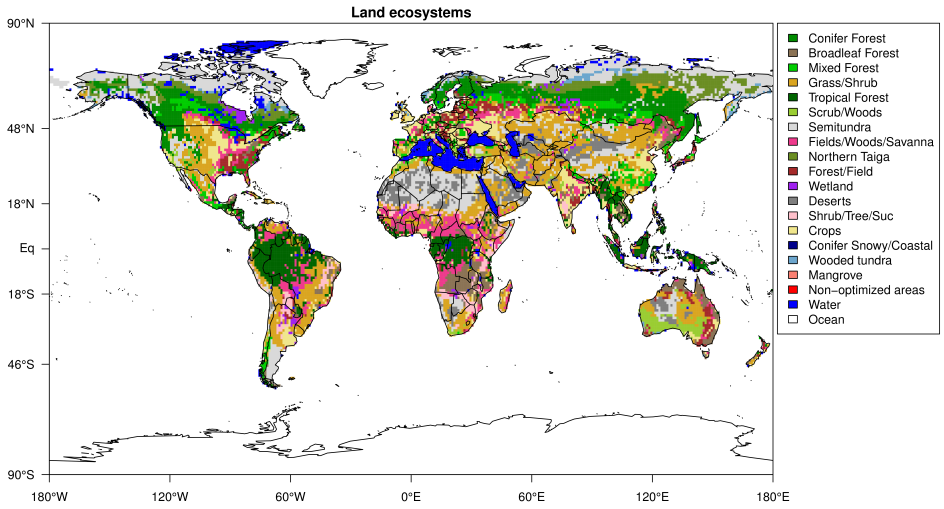


Figure C.1: Global distribution of Olson ecosystem types.

Ecosystem Type	North American Boreal		North American Temperate	
	Area (km^2)	Percentage	Area (km^2)	Percentage
Conifer Forest	2315376	22.9%	1607291	14.0%
Broadleaf Forest	–	–	269838	2.4%
Mixed Forest	592291	5.9%	930813	8.1%
Grass/Shrub	53082	0.5%	2515582	21.9%
Tropical Forest	–	–	58401	0.5%
Scrub/Woods	–	–	416520	3.6%
Semitundra	3396292	33.6%	866468	7.6%
Fields/Woods/Savanna	29243	0.3%	1020939	8.9%
Northern Taiga	1658773	16.4%	–	–
Forest/Field	61882	0.6%	1243174	10.8%
Wetland	322485	3.2%	66968	0.6%
Deserts	–	–	21934	0.2%
Shrub/Tree/Suc	–	–	11339	0.1%
Crops	–	–	1969912	17.2%
Conifer Snowy/Coastal	41440	0.4%	73437	0.6%
Wooded tundra	360388	3.6%	6643	0.1%
Mangrove	–	–	–	–
Non-optimized areas	–	–	–	–
Water	1269485	12.6%	384728	3.4%
Total	10100736	100.0%	11463986	100.0%

Table C.1: Ecosystem areas over the two Transcom regions covering North America.

C.2 Why use ecoregions?

A fundamental challenge to atmospheric inversions like CarbonTracker is that there are not enough observations to directly constrain fluxes at all times and in all places. It is therefore necessary to find a way to reduce the number of unknowns being estimated. Strategies to reduce the number of unknowns in problems like this one generally impose information from external sources. In CarbonTracker, we reduce the problem size both by estimating fluxes at the ecoregion scale, and by using a terrestrial biological model to give a first guess flux from the ecoregion. The model is also used to give the spatial and temporal distribution of CO_2 flux within a region and week.

C.3 Ecosystems within Transcom regions

Each Transcom land region (Figure C.2) can contain up to 19 ecoregions.

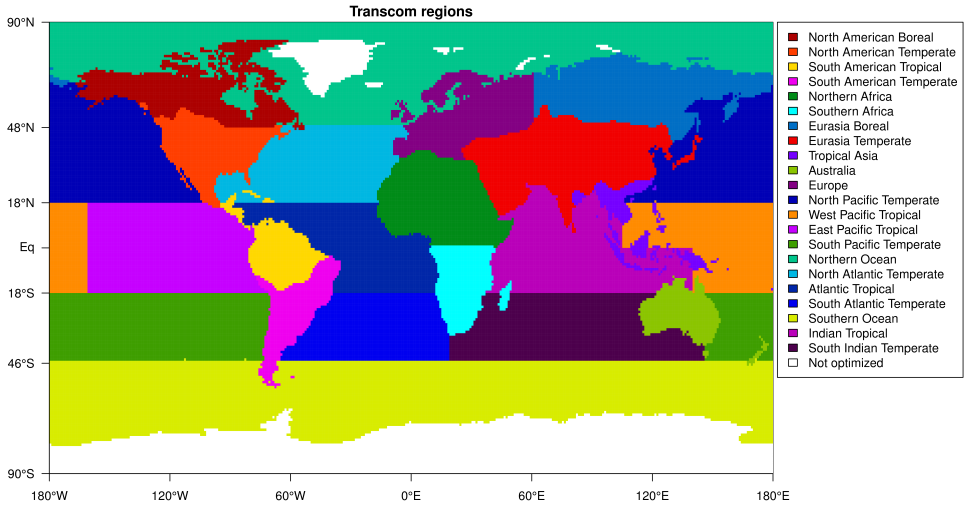


Figure C.2: The 11 land regions and 11 ocean regions of the Transcom project, along with the unoptimized land areas in Greenland and Antarctica

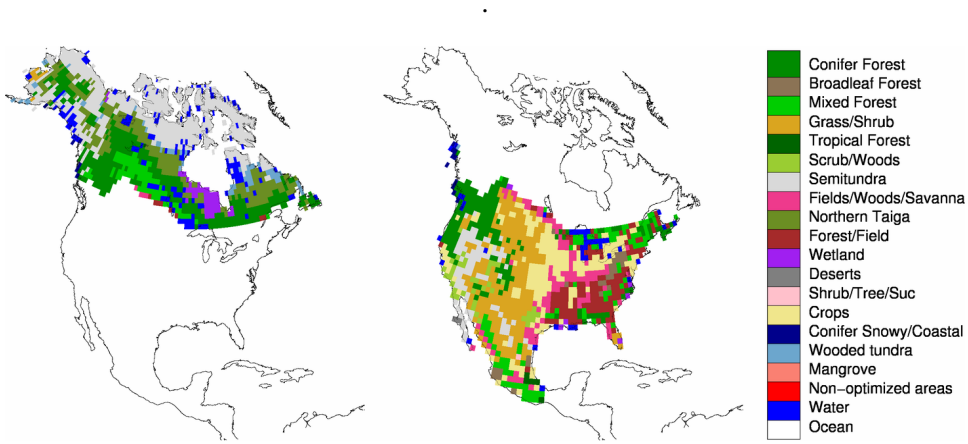


Figure C.3: Ecoregions within the North American Boreal (left) and North American Temperate (right) Transcom regions.

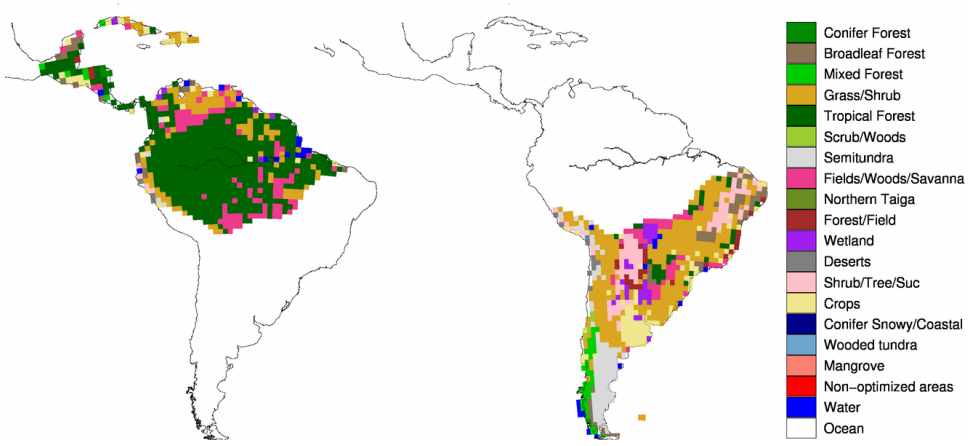


Figure C.4: Ecoregions within the South American Tropical (left) and South American Temperate (right) Transcom regions.

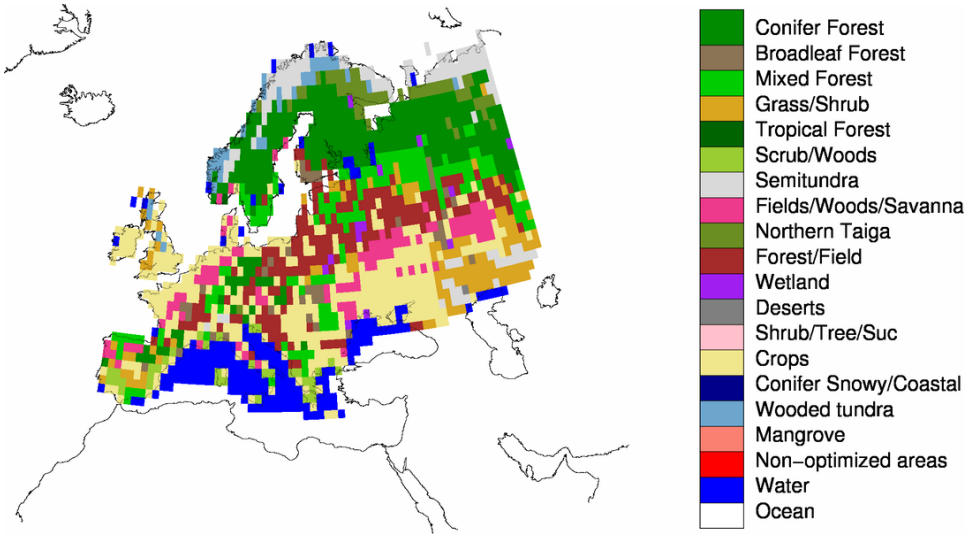


Figure C.5: Ecoregions within the Europe Transcom region.

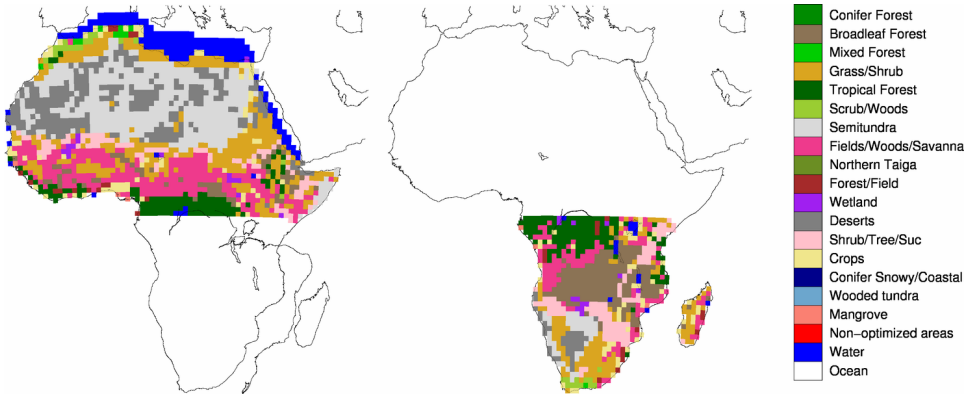


Figure C.6: Ecoregions within the Northern Africa (left) and Southern Africa (right) Transcom regions.