3-7-2023

National CO₂ budgets (2015-2020) inferred from atmospheric CO₂ observations in support of the global stocktake

Brendan Byrne
_California Institute of Technology_

David F. Baker
_Colorado State University_

Sourish Basu
_NASA Goddard Space Flight Center_

Michael Bertolacci
_University of Wollongong_

Kevin W. Bowman
_California Institute of Technology_

See next page for additional authors

Follow this and additional works at: https://scholarworks.sjsu.edu/faculty_rsca

Recommended Citation

Brendan Byrne; David F. Baker; Sourish Basu; Michael Bertolacci; Kevin W. Bowman; Dustin Carroll; Abhishek Chatterjee; Frédéric Chevallier; Philippe Ciais; Noel Cressie; David Crisp; Sean Crowell; Feng Deng; Zhu Deng; Nicholas M. Deutscher; Manvendra K. Dubey; Sha Feng; Omaira E. García; and For full author list, see comments below. "National CO₂ budgets (2015-2020) inferred from atmospheric CO₂ observations in support of the global stocktake" _Earth System Science Data_ (2023): 963-1004.
https://doi.org/10.5194/essd-15-963-2023

This Article is brought to you for free and open access by SJSU ScholarWorks. It has been accepted for inclusion in Faculty Research, Scholarly, and Creative Activity by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.
Authors
Brendan Byrne; David F. Baker; Sourish Basu; Michael Bertolacci; Kevin W. Bowman; Dustin Carroll; Abhishek Chatterjee; Frédéric Chevallier; Philippe Ciais; Noel Cressie; David Crisp; Sean Crowell; Feng Deng; Zhu Deng; Nicholas M. Deutscher; Manvendra K. Dubey; Sha Feng; Omaira E. García; and For full author list, see comments below

This article is available at SJSU ScholarWorks: https://scholarworks.sjsu.edu/faculty_rsca/3193
National CO₂ budgets (2015–2020) inferred from atmospheric CO₂ observations in support of the global stocktake

Brendan Byrne¹, David F. Baker², Sourish Basu³, Michael Bertolacci⁵, Kevin W. Bowman¹,⁶, Dustin Carroll⁷, Abhishek Chatterjee⁷, Frédéric Chevallier⁸, Philippe Ciais⁸, Noel Cressie⁵,¹, David Crisp⁹, Sean Crowell¹⁰, Feng Deng¹⁰, Zhu Deng¹¹, Nicholas M. Deutscher¹², Manvendra K. Dubey¹³, Sha Feng¹⁴, Omaira E. Garcia¹⁵, David W. T. Griffith¹², Benedikt Herkommer¹⁶, Lei Hu¹⁷, Andrew R. Jacobson¹⁷,¹⁸, Rajesh Janardanan¹⁹, Sujong Jeong²⁰, Matthew S. Johnson¹¹, Dylan B. A. Jones¹⁰, Rigel Kivi²², Junjie Liu¹,²³, Zhiqiang Liu²⁴, Shamil Maksyutov¹⁹, John B. Miller¹⁷, Scot M. Miller²⁵, Isamu Morino¹⁹, Justus Notholt²⁶, Tomohiro Oda²⁷,²⁸, Christopher W. O’Dell², Young-Suk Oh²⁹, Hirofumi Ohyama¹⁹, Prabir K. Patra³⁰, Hélène Peiro⁹, Christof Petri²⁶, Sajeev Philip³¹, David F. Pollard³², Benjamin Poulter³, Marine Remaud⁸, Andrew Schuh², Mahesh K. Sha³³, Kei Shiotomi³⁴, Kimberly Strong¹⁰, Colm Sweeney¹⁷, Yao Te³⁵, Hanqin Tian³⁶,³⁷, Voltaire A. Velasco¹²,³⁸, Mihalis Vrekoussis³⁹,²⁶, Thorsten Warneke³⁶, John R. Worden¹, Debra Wunch¹⁰, Yuanzhi Yao³⁶, Jeongmin Yun³⁰, Andrew Zammit-Mangion³, and Ning Zeng²⁸,⁴

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
²Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, CO, USA
³NASA Goddard Space Flight Center, Global Modeling and Assimilation Office, Greenbelt, MD, USA
⁴Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
⁵School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, Australia
⁶Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA
⁷Moss Landing Marine Laboratories, San José State University, Moss Landing, CA, USA
⁸Laboratoire des Sciences du Climat et de L’Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, 91191 Gif-sur-Yvette, France
⁹GeoCarb Mission Collaboration, University of Oklahoma, Norman, OK, USA
¹⁰Department of Physics, University of Toronto, Toronto, Ontario, Canada
¹¹Department of Earth System Science, Tsinghua University, Beijing, China
¹²Centre for Atmospheric Chemistry, School of Earth, Atmospheric and Life Sciences, University of Wollongong, Wollongong, NSW, Australia
¹³Earth System Observation, Los Alamos National Laboratory, Los Alamos, NM, USA
¹⁴Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA
¹⁵Izaña Atmospheric Research Center (IARC), State Meteorological Agency of Spain (AEMet), Tenerife, Spain
¹⁶Institute for Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
¹⁷NOAA Global Monitoring Laboratory, Boulder, CO, USA
¹⁸Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA
¹⁹Satellite Observation Center, Earth System Division, National Institute for Environmental Studies, Tsukuba, Japan
²⁰Department of Environmental Planning, Graduate School of Environmental Studies, Seoul National University, Seoul, Republic of Korea

Published by Copernicus Publications.
Abstract. Accurate accounting of emissions and removals of CO$_2$ is critical for the planning and verification of emission reduction targets in support of the Paris Agreement. Here, we present a pilot dataset of country-specific net carbon exchange (NCE; fossil plus terrestrial ecosystem fluxes) and terrestrial carbon stock changes aimed at informing countries’ carbon budgets. These estimates are based on “top-down” NCE outputs from the v10 Orbiting Carbon Observatory (OCO-2) modeling intercomparison project (MIP), wherein an ensemble of inverse modeling groups conducted standardized experiments assimilating OCO-2 column-averaged dry-air mole fraction (X$_{CO_2}$) retrievals (ACOS v10), in situ CO$_2$ measurements or combinations of these data. The v10 OCO-2 MIP NCE estimates are combined with “bottom-up” estimates of fossil fuel emissions and lateral carbon fluxes to estimate changes in terrestrial carbon stocks, which are impacted by anthropogenic and natural drivers. These flux and stock change estimates are reported annually (2015–2020) as both a global 1° × 1° gridded dataset and a country-level dataset and are available for download from the Committee on Earth Observation Satellites’ (CEOS) website: https://doi.org/10.48588/npf6-sw92 (Byrne et al., 2022). Across the v10 OCO-2 MIP experiments, we obtain increases in the ensemble median terrestrial carbon stocks of 3.29–4.58 Pg CO$_2$ yr$^{-1}$ (0.90–1.25 Pg C yr$^{-1}$). This is a result of broad increases in terrestrial carbon stocks across the northern extratropics, while the tropics generally have stock losses but with considerable regional variability and differences between v10 OCO-2 MIP experiments. We discuss the state of the science for tracking emissions and removals using top-down methods, including current limitations and future developments towards top-down monitoring and verification systems.
1 Introduction

To reduce the risks and impacts of climate change, the Paris Agreement aims to limit the global average temperature increase to well below 2 °C above pre-industrial levels and to pursue efforts to limit these increases to less than 1.5 °C. To this end, each Party to the Paris Agreement agreed to prepare and communicate successive nationally determined contributions (NDCs) of greenhouse gas (GHG) emission reductions. Collective progress toward this goal of the Paris Agreement is evaluated in global stocktakes (GSTs), which are conducted at 5-year intervals; the first GST is scheduled in 2023. The outcome of each GST is then used as input, or as a “ratchet mechanism”, for new NDCs that are meant to encourage greater ambition.

In support of the first GST, Parties to the Paris Agreement are compiling national GHG inventories (NGHGIs) of emissions and removals, which are submitted to the United Nations Framework Convention of Climate Change (UNFCCC) and inform their progress toward the emission-reduction targets in their individual NDCs. For these inventories, emissions and removals are generally estimated using “bottom-up” approaches, wherein CO₂ emission estimates are based on activity data and emission factors, while CO₂ removals by sinks are based on inventories of carbon stock changes and models, following the methods specified in the 2006 Intergovernmental Panel on Climate Change (IPCC) Guidelines for National GHG Inventories (IPCC, 2006). This approach allows for explicit characterization of CO₂ emissions and removals into five categories: energy; industrial processes and product use (IPPU); agriculture; land use, land-use change and forestry (LULUCF); and waste. Bottom-up methods can provide precise and accurate country-level emission estimates when the activity data and emission factors are well quantified and understood (Petrescu et al., 2021), such as for the fossil fuel combustion category of the energy sector in many countries. However, these estimates can have considerable uncertainty when the emission processes are challenging to quantify (such as for agriculture, LULUCF and waste) or if the activity data are inaccurate or missing. For example, Grassi et al. (2022) and McGlynn et al. (2022) estimate the uncertainty on the net LULUCF CO₂ flux to be roughly 35 % for Annex I countries and 50 % for non-Annex I countries. In addition, these estimates do not capture carbon emissions and removals from unmanaged systems, which are not directly considered in the Paris Agreement, but impact the global carbon budget and growth rate of atmospheric CO₂.

As a complement to these accounting-based inventory efforts, an independent “top-down” assessment of net surface-atmosphere CO₂ fluxes may be obtained from ground-based, airborne and space-based observations of atmospheric CO₂ mole fractions. These top-down methods have undergone rapid improvements in recent years, as recognized in the 2019 Refinement to the 2006 IPCC Guidelines for National GHG Inventories (IPCC, 2019). And, although these methods were not deemed to be a standard tool for verification of conventional inventories, a number of countries (UK, Switzerland, USA and New Zealand) have adopted atmospheric inverse modeling as a verification system in national inventory reports. Initially, these countries have focused on non-CO₂ gasses (e.g., EPA, 2022), but top-down assessments of the CO₂ budget are now under development in New Zealand (https://niwa.co.nz/climate/research-projects/carbon-watch-nz, last access: 6 February 2023). Furthermore, significant investments towards building anthropogenic CO₂ emissions monitoring and verification support capacity are ongoing within the European Commission’s Copernicus Program (see Sect. 9.2.1).

In top-down CO₂ flux estimation, the net surface-atmosphere CO₂ fluxes are inferred from atmospheric CO₂ observations using state-of-the-art atmospheric CO₂ inversion systems (e.g., Peiro et al., 2022). This approach provides spatially and temporally resolved estimates of surface-atmosphere fluxes for land and ocean regions from which country-level annual land–atmosphere CO₂ fluxes can be estimated. The impact of fossil fuel (and usually fire CO₂ emissions) on the observations is accounted for in the inversions by prescribing maps of those emissions and assuming that they are perfectly known. Thus, fossil fuel and fire CO₂ emissions are not diagnosed yet by these inversions but net surface–atmosphere CO₂ fluxes from the terrestrial biosphere and oceans are. Terrestrial carbon stock changes can then be calculated by combining net surface–atmosphere CO₂ fluxes with estimates of fossil fuel emissions and horizontal (“lateral”) fluxes occurring within the terrestrial biosphere or between the land and ocean (Kondo et al., 2020). One example of a lateral flux is harvested agricultural products, where carbon is sequestered from the atmosphere by photosynthesis in one region, but then this carbon is harvested and exported to another region as agricultural products. Similarly, carbon sequestered by photosynthesis in a forest can be leached away by streams and rivers and then exported to the ocean. These lateral carbon fluxes are not directly identifiable in atmospheric CO₂ measurements, but accounting for their impact is required in order to convert net land fluxes into stock changes. These estimated terrestrial carbon stock changes reflect the combined impact of direct anthropogenic activities and changes to both managed and unmanaged ecosystems in response to rising CO₂, climate change and disturbance events (such as fires).

The top-down budgets presented here extend several previous studies that have developed approaches to compare inversion results to NGHGIs. Ciais et al. (2021) proposed a protocol for reporting bottom-up and top-down fluxes so that they can be compared consistently. Petrescu et al. (2021) compared top-down fluxes with inventory estimates for the
European Union and UK, including for an ensemble of regional inversions over Europe (Monteil et al., 2020). Chevalier (2021) noted that inversion results for terrestrial CO₂ fluxes should be restricted to managed lands and applied a managed land mask to the gridded fluxes of the Copernicus Atmosphere Monitoring Service (CAMS) CO₂ inversions for the comparison to UNFCCC values in 10 large countries or groups of countries. Deng et al. (2022) compared CO₂, CH₄ and N₂O fluxes from inversion ensembles available from the Global Carbon Project. For CO₂, they used six CO₂ flux estimates from inverse models that assimilated measurements from the global air-sample network, filtered their results over managed lands and corrected them for CO₂ fluxes induced by lateral processes to compare with carbon stock changes reported to the UNFCCC by a set of 12 countries. We expand upon these previous studies by providing top-down CO₂ budgets from the v10 Orbiting Carbon Observatory Model Intercomparison Project (v10 OCO-2 MIP), wherein an ensemble of inverse modeling groups conducted standardized experiments assimilating OCO-2 column-averaged dry-air mole fraction (XCO₂) retrievals (retrieved with version 10 of the Atmospheric CO₂ Observations from Space (ACOS) full-physics retrieval algorithm), in situ CO₂ measurements or combinations of these data. This allows us to quantify the sensitivity of top-down carbon budget estimates to the inversion modeling system and the atmospheric CO₂ dataset used to constrain flux estimates.

This paper is outlined as follows. The remainder of Sect. 1 describes the objectives of this work (Sect. 1.1) and provides background information on both the global carbon cycle (Sect. 1.2) and top-down atmospheric CO₂ inversions (Sect. 1.3). Section 2 defines the carbon cycle fluxes of interest. Section 3 describes the flux datasets and their uncertainties, including fossil fuel emissions, the v10 OCO-2 MIP, riverine fluxes, wood fluxes, crop fluxes and the net terrestrial carbon stock loss. Section 4 provides an evaluation of the v10 OCO-2 MIP flux estimates. Section 5 presents two metrics for interpreting the top-down constraints on the CO₂ budget. Section 6 gives a description of the dataset. Sect. 7 shows the characteristics of the dataset. Sect. 8 demonstrates how these data can be compared with national inventories, and Sect. 9 discusses current limitations and future directions. Section 10 describes the data availability. Finally, Sect. 11 gives the conclusions of this study.

1.1 Objectives

This is a pilot project designed to start a dialogue between the top-down research community, inventory compilers and the GHG assessment community to identify ways that top-down CO₂ flux estimates can help inform country-level carbon budgets (see Worden et al., 2022, for a similar pilot methane dataset). To meet this objective, the primary goal of this work is to provide two products: (1) annual net surface–atmosphere CO₂ fluxes and (2) annual changes in terrestrial carbon stocks. These products are provided annually over the 6-year period 2015–2020 both on a 1° × 1° global grid and as country-level totals with error characterization.

These products are intended to be used to help inform inventory development and identify areas for future research in both top-down and bottom-up approaches, including informing strategies for operational top-down carbon cycle products that can be used for tracking combined changes in managed and unmanaged carbon stocks and that can help quantify the impact of emission reduction activities.

1.2 Overview of the carbon cycle

The burning of fossil fuels and cement production release geologic carbon to the atmosphere (40.0 ± 3.3 Pg CO₂ yr⁻¹ or 10.9 ± 0.9 Pg C yr⁻¹ over 2010–2019; Canadell et al., 2021). These emissions, along with land-use activities, impact carbon cycling between atmospheric, oceanic and biospheric reservoirs that make up a near-closed system on annual timescales. As a result, roughly half of the emitted CO₂ from anthropogenic sources is absorbed by terrestrial ecosystems and oceans (Friedlingstein et al., 2022), reducing the rate of atmospheric CO₂ increase (18.7 ± 0.08 Pg CO₂ yr⁻¹ or 5.1 ± 0.02 Pg C yr⁻¹ over 2010–2019; Canadell et al., 2021). Here we briefly review the movement of carbon between the reservoirs and how these processes are modulated by human activities.

Fluxes of carbon between the atmosphere and ocean are driven by the difference in partial pressures of CO₂ between seawater and air, resulting in roughly balancing fluxes from the ocean-to-atmosphere and atmosphere-to-ocean of ~293 Pg CO₂ yr⁻¹ (~80 Pg C yr⁻¹) each way (Ciais et al., 2013), with a residual net atmospheric-to-ocean flux due to increasing atmospheric CO₂ (9.2 ± 2.2 Pg CO₂ yr⁻¹ or 2.5 ± 0.6 Pg C yr⁻¹ over 2010–2019; Canadell et al., 2021). Regional variations in the solubility and saturation of CO₂ in ocean waters drive net fluxes, with net fluxes to the atmosphere in upwelling regions, such as the eastern boundary of basins and in equatorial zones (McKinley et al., 2017). Meanwhile, there are net removals by the ocean in western boundary currents and at extratropical latitudes (McKinley et al., 2017). Within the oceans, circulation patterns, mixing and biologic activity act to redistribute carbon.

On land, terrestrial ecosystems remove atmospheric carbon through photosynthesis, referred to as gross primary production (GPP) (Fig. 1). GPP draws roughly 440 Pg CO₂ yr⁻¹ (120 Pg C yr⁻¹) from the atmosphere (Anav et al., 2015). Roughly half of this carbon is emitted back to the atmosphere by plants through autotrophic respiration, while the remaining carbon is used to generate plant biomass and is referred to as net primary production (NPP). On an annual basis, the carbon sequestered through NPP is roughly balanced by carbon loss through a number of processes. The largest of these processes is heterotrophic respiration, which is the respiratory emission of CO₂ (from the dead organic
matter and soil carbon pools) by heterotrophic organisms, and accounts for 82–95 % of NPP (Randerson et al., 2002). The combination of heterotrophic and autotrophic respiration is called ecosystem respiration ($R_{ECO}$). The remaining processes have smaller magnitudes but are still critical for determining the carbon balance of ecosystems. Biomass burning, the emission of carbon to the atmosphere through combustion, releases roughly 7.3 Pg CO$_2$ yr$^{-1}$ (2 Pg C yr$^{-1}$) to the atmosphere on an annual basis but with considerable interannual variability (van der Werf et al., 2017). Carbon can also be emitted from the terrestrial biosphere to the atmosphere in the form of carbon monoxide (CO), methane (CH$_4$) and other biologic volatile organic compounds (BVOCs), which are oxidized to CO$_2$ in the atmosphere. Rivers move carbon in the form of dissolved inorganic carbon (DIC), dissolved organic carbon (DOC) and particulate organic carbon (POC). This carbon of terrestrial origin is partly transported to the open ocean, partly released to the atmosphere from inland waters and estuaries, and partly buried in aquatic or marine sediments. Finally, anthropogenic activities such as harvesting of crop and wood products result in lateral transport of carbon such that the removal of atmospheric CO$_2$ through NPP and emission of atmospheric CO$_2$ through respiration (e.g., decomposition in a landfill) or combustion (e.g., burning of biofuels) occur in different regions. See Fig. 1 for an illustration of these fluxes.

Globally, there is a long-term net uptake of atmospheric CO$_2$ by the land (approximately $-6.6$ Pg CO$_2$ yr$^{-1}$ or $-1.8$ Pg C yr$^{-1}$ over 2010–2019; Canadell et al., 2021), which is the residual of an emission due to net land-use change $(5.9 \pm 2.6$ Pg CO$_2$ yr$^{-1}$ or $1.6 \pm 0.7$ Pg C yr$^{-1}$ over 2010–2019; Canadell et al., 2021) and removal by other terrestrial ecosystems $(12.6 \pm 3.3$ Pg CO$_2$ yr$^{-1}$ or $3.4 \pm 0.9$ Pg C yr$^{-1}$ over 2010–2019; Canadell et al., 2021). This removal is partially driven by direct feedbacks between increasing CO$_2$ and the biosphere, such as CO$_2$ fertilization of photosynthesis and increased water use efficiency. Carbon-climate feedbacks also lead to both increases and decreases in terrestrial carbon stocks: for example, warming at high latitudes leads to a more productive biosphere, but it also leads to increased plant and soil respiration (Kaushik et al., 2020; Walker et al., 2021; Canadell et al., 2021; Crisp et al., 2022). In addition, the release of nitrogen through anthropogenic energy and fertilizer use may drive increased carbon sequestration by the terrestrial biosphere (Schulte-Uebbing et al., 2022; Y. Liu et al., 2022; Lu et al., 2021). Regrowth of forests in previously cleared areas, especially in the extratropics, is also thought to be an important uptake term (Kondo et al., 2018; Cook-Patton et al., 2020). Currently, the relative impact of each of these contributions to long-term terrestrial carbon sequestration is poorly known and likely varies between biomes and climates.

While the existence of a long-term global land sink is supported through a number of lines of evidence (Ballantyne et al., 2012; Keeling and Graven, 2021), regional-scale emissions and removals are less well quantified. Regional-scale carbon sequestration can differ substantially from the global mean and can be impacted by the regional climate, disturbance events (Frank et al., 2015; Wang et al., 2021) and anthropogenic activities (Caspersen et al., 2000; Harris et al., 2012). The need to better quantify regional-scale emissions and removals of carbon has motivated much of the recent expansion of in situ CO$_2$ observing networks, the launch of space-based CO$_2$ observing systems and the development of CO$_2$ inversion systems.

1.3 Background on atmospheric CO$_2$ inversions

Atmospheric CO$_2$ inversions estimate the underlying net surface–atmosphere CO$_2$ fluxes from atmospheric CO$_2$ observations, and this is what is meant by the top-down approach (Bolin and Keeling, 1963; Tans et al., 1990; Enting et al., 1995; Gurney et al., 2002; Peiro et al., 2022). In this approach, an atmospheric chemical transport model (CTM) is employed to relate surface–atmosphere CO$_2$ fluxes to observed atmospheric CO$_2$ mole fractions. As an inverse problem, the upwind CO$_2$ fluxes are estimated from the downwind observed CO$_2$ mole fractions. The surface CO$_2$ fluxes are adjusted so that forward-simulated CO$_2$ mole fractions better match the CO$_2$ measurements while considering the uncertainty statistics on the observations, transport and prior surface fluxes.

The atmospheric CO$_2$ inversion problem is generally ill-posed such that the solution is underdetermined by the observational constraints. In this case, additional information is required to produce a unique solution and prevent overfitting of the data (Lawson and Hanson, 1995; Tarantola, 2005). Typically, this is performed using Bayesian inference, where prior mean fluxes and their uncertainties provide additional information required to estimate fluxes (Rayner et al., 2019). Prior mean fluxes of net ecosystem exchange are usually obtained from terrestrial biosphere models (such as CASA, ORCHIDEE and CARDAMOM), while prior mean air–sea fluxes are derived from surface water partial pressure of CO$_2$ ($p$CO$_2$) datasets or from ocean models (e.g., Petro et al., 2022). The resulting posterior flux estimates combine the constraints on surface fluxes from atmospheric CO$_2$ data with the prior knowledge of the fluxes. If there is a high density of assimilated CO$_2$ observations, then the posterior fluxes will be more strongly impacted by the assimilated data, whereas, in regions with sparse observational coverage, the posterior fluxes will generally remain similar to the prior fluxes (assuming similar prior flux uncertainties across regions).

Measurements of atmospheric CO$_2$ best inform diffuse biosphere–atmosphere fluxes on large spatial scales. This is because CO$_2$ has a long atmospheric lifetime such that the perturbation to atmospheric CO$_2$ due to emissions and removals from individual processes and locations gets mixed in the atmosphere (Gloor et al., 2001; Liu et al., 2015). For
Figure 1. CO$_2$ is removed from the atmosphere through photosynthesis (GPP) and then emitted back to the atmosphere through a number of processes. Three processes move carbon laterally on Earth’s surface such that emissions of CO$_2$ occur in a different region than removals. (1) Agriculture: harvested crops are transported to urban areas and to livestock, which are themselves exported to urban areas. CO$_2$ is respired to the atmosphere in livestock or urban areas. (2) Forestry: logged carbon is transported to urban and industrial areas, then emitted through decomposition in a landfill or combustion as a biofuel. (3) Water cycle: carbon is leached from soils into water bodies, such as lakes. The carbon is then either deposited, released to the atmosphere or transported to the ocean (Regnier et al., 2022). Arrows show carbon fluxes, and colors indicate whether the flux is associated with (grey) fossil fuel emissions, (dark green) ecosystem metabolism, (red) biomass burning, (light green) forestry, (yellow) agriculture or (blue) the water cycle. Semi-transparent arrows show fluxes that move between the surface and atmosphere, while solid arrows show fluxes that move between land regions. Dashed arrows show surface–atmosphere fluxes of reduced carbon species that are oxidized to CO$_2$ in the atmosphere. For simplicity, a cement carbonation sink, volcano emissions and a weathering sink are not included in this figure.

Example, the measurements of CO$_2$ at Mauna Loa, Hawaii, provide a good estimate of the global-scale changes of CO$_2$ surface fluxes. Inferring smaller-scale flux signals requires a high density of CO$_2$ observations (to capture gradients in atmospheric CO$_2$) and accurate modeling of atmospheric transport (to relate the measurements with surface fluxes). The accuracy of flux estimates depends on a number of factors, particularly the accuracy and precision of the data, transport model and prior constraints. Stringent requirements on the accuracy of space-based column-averaged dry-air mole fraction ($X_{CO_2}$) retrievals are required to infer surface fluxes (Chevallier et al., 2005a; Miller et al., 2007). Biases in $X_{CO_2}$ retrievals from the Orbiting Carbon Observatory (OCO-2) related to spectroscopic errors, solar zenith angle, surface properties, and atmospheric scattering by clouds and aerosols have been identified (Wunch et al., 2017b). However, intensive research has reduced retrieval errors over time (O’Dell et al., 2018; Kiel et al., 2019). As will be shown in Sect. 4.1, biases in OCO-2 $X_{CO_2}$ retrievals over land are thought to be relatively small, although regionally structured biases may be present. However, OCO-2 $X_{CO_2}$ retrievals over oceans may contain more large-scale spatially coherent retrieval errors that can adversely impact flux estimates.

Accurate atmospheric transport is critical for correctly relating surface–atmosphere fluxes to observations. Due to computational constraints, CTMs are typically run offline with coarsened meteorological fields relative to the parent numerical weather prediction model, which has been shown to introduce systematic transport errors in some configurations (Yu et al., 2018; Stanovich et al., 2020). In addition, these offline CTMs have been shown to have large-scale systematic differences in transport associated with the implementation of transport algorithms (Schuh et al., 2019, 2022). These errors appear to be of the same order as the retrieval biases, although the patterns in time and space are different. Systematic errors related to model transport (and errors in prior information) can partially be accounted for by performing multiple inversions that differ in CTM and prior constraints employed. This motivates inversion model intercomparison projects (MIPs), such as the OCO-2 MIP project.
(see Sect. 3.2; Crowell et al., 2019; Peiro et al., 2022). From these ensembles of inversions, estimates of both systematic errors (accuracy) and random errors (precision) can be obtained from the model spread.

2 Definitions

In this work, we focus on the carbon budget of Earth’s land area, including aquatic systems such as rivers and lakes. In particular, we consider fluxes of carbon between the land and the atmosphere and lateral carbon transport processes on land and between the land and ocean (Fig. 1). We define the following annual net carbon fluxes (see Fig. 2 for a schematic representation of these fluxes):

- **Fossil fuel and cement emissions (FF).** The burning of fossil fuels and release of carbon due to cement production, representing a flux of carbon from the land surface (geologic reservoir) to the atmosphere.

- **Net biosphere exchange (NBE).** Net flux of carbon from the terrestrial biosphere to the atmosphere due to biomass burning (BB) and \( R_{\text{eco}} \) minus gross primary production (GPP) (i.e., \( \text{NBE} = \text{BB} + R_{\text{eco}} - \text{GPP} \)). It includes both anthropogenic processes (e.g., deforestation, reforestation, farming) and natural processes (e.g., climate-variability-induced carbon fluxes, disturbances, recovery from disturbances).

- **Terrestrial net carbon exchange (NCE).** Net flux of carbon from the surface to the atmosphere. For land, NCE can be defined as

\[
\text{NCE} = \text{NBE} + \text{FF}.
\]

- **Lateral crop flux (\( F_{\text{crop trade}} \)).** The lateral flux of carbon in (positive) or out (negative) of a region due to agriculture.

- **Lateral wood flux (\( F_{\text{wood trade}} \)).** The lateral flux of carbon in (positive) or out (negative) of a region due to wood product harvesting and usage.

- **Lateral river flux (\( F_{\text{rivers export}} \)).** The lateral flux of carbon in (positive) or out (negative) of a region transported by the water cycle.

- **Net terrestrial carbon stock loss (\( \Delta C_{\text{loss}} \)).** Positive values indicate a loss (decrease) of terrestrial carbon stocks, and this is the negative of \( \Delta C_{\text{gain}} \):

\[
\Delta C_{\text{gain}} = -\Delta C_{\text{loss}}.
\]

### Country and regional aggregation

To aggregate gridded \( 1^\circ \times 1^\circ \) flux estimates to country totals we use a country mask (Center for International Earth Science Information Network – CIESIN – Columbia University, 2018). We also provide NCE and \( \Delta C_{\text{loss}} \) estimates for several country groupings. A number of regional intergovernmental organizations are included: the Association of Southeast Asian Nations (ASEAN), the African Union (AU) and each of its sub-regions (North, South, West, East and Central), the Community of Latin American and Caribbean States plus Brazil (CELAC+Brazil), the Economic Cooperation Organization (ECO), the European Union (EU or EU27), and the South Asian Association for Regional Cooperation (SAARC). We also include some geographic regions, specifically North America, the Middle East and Europe. Countries included in these groupings are listed in the Supplement (Text S1).

### 3 Flux datasets

Here, we describe the methodologies and datasets for estimating FF (Sect. 3.1), NCE (Sect. 3.2) and lateral carbon fluxes (Sect. 3.3), as well as how these data are used to estimate \( \Delta C_{\text{loss}} \) (Sect. 3.4).

#### 3.1 Fossil fuel and cement emissions

Gridded \( 1^\circ \times 1^\circ \) fossil CO\(_2\) emissions, including those from cement production, are calculated as follows. Monthly gridded emissions up to 2019 are taken from the 2020 version of the Open-source Data Inventory for Anthropogenic CO\(_2\) (ODIAC2020, 2000–2019) emission data product (Oda and Maksyutov, 2011; Oda et al., 2018). The 2020 emissions were not part of ODIAC but were projected using the Carbon Monitor (CM) emission data product (https://carbonmonitor.org/, last access: 19 May 2021). For each month in 2020 and later, the ratio between that month’s emissions and the emissions from the same month in 2019 was calculated from the CM emission data. Since CM provides daily emissions per sector for a handful of major emitting countries and the globe, CM emissions are summed over sectors and days in each month to create monthly total emissions per named country and the rest of the world (RoW). The ratio of each (post-2019) month’s emission to the same month in 2019 is then calculated per named country and RoW, then distributed over a \( 1^\circ \times 1^\circ \) grid assuming homogeneity of the ratio over each named country and RoW. The 2019 ODIAC emissions...
Figure 2. Carbon fluxes for a given land region, such as a country. Boxes with solid backgrounds show reservoirs of carbon. Arrows with hatched shading show fluxes between reservoirs. NCE is underlined to emphasize that this quantity is estimated from the atmospheric CO$_2$ measurements using top-down methods. Italicized quantities are obtained from bottom-up datasets (FF, $F_{\text{crop trade}}$, $F_{\text{wood trade}}$, $F_{\text{rivers export}}$). Bold quantities are derived in this study from the top-down and bottom-up datasets (NBE, $\Delta C_{\text{gain}}$, $\Delta C_{\text{loss}}$).

for that month are then multiplied by the ratio to generate $1^\circ \times 1^\circ$ monthly emissions after 2019. While this method loses the information of day-to-day variability provided by CM, this is a conscious choice to be consistent over the entire inversion period. Finally, we impose day-of-week and hour-of-day variations on these fluxes following the Temporal Improvements for Modeling Emissions by Scaling (TIMES) diurnal and day-of-week scaling (Nassar et al., 2013). The $1^\circ \times 1^\circ$ uncertainty map is based on the combination of the global level FF uncertainty (1σ of 4.2 %, Andres et al., 2014) and the grid level emission differences due to the different disaggregation methods (Oda et al., 2015). Note that these FF uncertainties are not considered in the inversions used for this product development.

Country-level fossil fuel emission estimates are obtained by aggregating the $1^\circ \times 1^\circ$ estimates using the country mask. Uncertainties on country-level estimates are calculated using the fractional uncertainties of Andres et al. (2014).

3.2 Net carbon exchange (NCE) and net biosphere exchange (NBE)

We employ results from the v10 OCO-2 MIP, which is an international collaboration of atmospheric CO$_2$ inversion mod-
Table 1. Inversion specifications for each v10 OCO-2 MIP ensemble member. Note that since the creation of this table, the Global Carbon Assimilation System (GCAS; Jiang et al., 2021) has also contributed an ensemble member.

<table>
<thead>
<tr>
<th>Simulation name (reference)</th>
<th>Transport model</th>
<th>Driving meteorology</th>
<th>Meteorology resolution (°)</th>
<th>Prior NEE</th>
<th>Prior air–sea</th>
<th>Prior fire</th>
<th>Inverse method</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMES (Philip et al., 2019, 2022)</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>4° × 5°</td>
<td>CASA-GFED4.1s</td>
<td>CT2019OI</td>
<td>GFEDv4.1s</td>
<td>4D-Var</td>
</tr>
<tr>
<td>Baker (D. F. Baker et al., 2006; Baker et al., 2010)</td>
<td>PCTM</td>
<td>MERRA-2</td>
<td>1° × 1.25° prior, 4° × 5° opt</td>
<td>CASA-GFED3</td>
<td>Landschützer v4.4</td>
<td>GFEDv4</td>
<td>4D-Var</td>
</tr>
<tr>
<td>CAMS (Chevallier et al., 2005b)</td>
<td>LMDz</td>
<td>ERA5</td>
<td>1.9° × 3.75°</td>
<td>ORCHIDEE (climato-logical)</td>
<td>CMEMS</td>
<td>GFED4.1s</td>
<td>Variational</td>
</tr>
<tr>
<td>CMS-Fluxa (J. Liu et al., 2021)</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>4° × 5°</td>
<td>CARDAMOM</td>
<td>MOM-6</td>
<td>CARDAMOM</td>
<td>4D-Var</td>
</tr>
<tr>
<td>COLA (Z. Liu et al., 2022)</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>4° × 5°</td>
<td>VEGAS</td>
<td>Rödenbeck 2021</td>
<td>VEGAS</td>
<td>EnKF</td>
</tr>
<tr>
<td>CSU</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>4° × 5°</td>
<td>SiB-4 w/MERRA-2</td>
<td>Landschützer v18</td>
<td>GFED4.1s</td>
<td>synthesis</td>
</tr>
<tr>
<td>CT (Jacobson et al., 2020)</td>
<td>TM5</td>
<td>ERA5</td>
<td>2° × 3°/1° × 1°</td>
<td>CT2019 CASA GFEDv4.1s</td>
<td>CT2019OI</td>
<td>CT2019 CASA GFED4.1s</td>
<td>EnKF</td>
</tr>
<tr>
<td>JHUa (Miller et al., 2020) (Chen et al., 2021b, a)</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>4° × 5°</td>
<td>CASA-GFED 4.1s</td>
<td>Takahashi</td>
<td>GFED4.1s</td>
<td>geostatistical/4D-Var</td>
</tr>
<tr>
<td>LoFIb (Weir et al., 2021a, b)</td>
<td>GEOS GCM</td>
<td>GEOS CGM</td>
<td>0.5° × 0.625°</td>
<td>CASA-GFED3</td>
<td>LoFI Takahashi</td>
<td>QFED</td>
<td>n/a</td>
</tr>
<tr>
<td>NIES (Maksyutov et al., 2021)</td>
<td>NIES-TM/FLEXPART</td>
<td>ERA5/JRA-55</td>
<td>3.75° × 3.75°/0.1° × 0.1°</td>
<td>Zeng 2020</td>
<td>Landschützer 2020</td>
<td>GFAS</td>
<td>4D-Var</td>
</tr>
<tr>
<td>OU</td>
<td>TM5</td>
<td>ERA-Interim</td>
<td>4° × 6°</td>
<td>CASA-GFED3</td>
<td>Takahashi</td>
<td>GFEDv3</td>
<td>4D-Var</td>
</tr>
<tr>
<td>TM5-4DVar</td>
<td>TM5</td>
<td>ERA-Interim</td>
<td>2° × 3°</td>
<td>SiB-CASA</td>
<td>CT2019 Opt Chim</td>
<td>GFEDv4</td>
<td>4D-Var</td>
</tr>
<tr>
<td>UT (Deng et al., 2014, 2016)</td>
<td>GEOS-Chem</td>
<td>GEOS-FP</td>
<td>4° × 5°</td>
<td>BEPS</td>
<td>Takahashi</td>
<td>GFEDv4</td>
<td>4D-Var</td>
</tr>
<tr>
<td>WOMBAT (Zammit-Mangion et al., 2022)</td>
<td>GEOS-Chem</td>
<td>MERRA-2</td>
<td>2° × 2.5°</td>
<td>SiB-4 w/MERRA-2</td>
<td>Landschützer 2020</td>
<td>GFED4.1s</td>
<td>Synthesis with MCMC</td>
</tr>
</tbody>
</table>

a Due to time constraints, this ensemble member is not included in this product. b This is not a traditional inversion and is not included in this product. n/a: not applicable.
– IS assimilates in situ CO₂ mole fraction measurements from an international observational network.

– LNLG assimilates ACOS v10 land nadir and land glint total column dry-air mole fractions (X_{\text{CO₂}}) from OCO-2.

– LNLGIS assimilates both in situ and ACOS v10 OCO-2 land nadir and glint X_{\text{CO₂}} retrievals together.

– OG assimilates ACOS v10 OCO-2 ocean glint X_{\text{CO₂}} retrievals.

– LNLGOGIS assimilates all the above datasets together.

For each experiment, each inversion group imposes a common fossil fuel emission dataset identical to the one described in Sect. 3.1. All other prior flux estimates were chosen independently by each modeling group and are listed in Table 1. The inversions assimilate the standardized v10 OCO-2 and in situ data from 6 September 2014 through 31 March 2021 (see Sect 3.2.1), with the length of spin-up period and in situ data assimilated during that period being left up to the discretion of each group in the MIP. Each modeling group submitted net air–sea fluxes and NBE across 2015–2020, interpolated from the native resolution to a 1° × 1° spatial grid at monthly resolution, which are publicly available for download from https://gml.noaa.gov/ccgg/OCO2_v10mip/ (last access: 6 February 2023).

The performance of each atmospheric CO₂ inversion was evaluated through comparisons of the posterior CO₂ mole-fraction field (i.e., CO₂ fields simulated forward with the posterior fluxes) against independent in situ CO₂ measurements and OCO-2 X_{\text{CO₂}} retrievals that were withheld from the assimilation for validation, as well as X_{\text{CO₂}} retrievals from the Total Column Carbon Observing Network (TCCON; Wunch et al., 2011). The evaluation of the experiments is presented in Sect. 4, with additional analysis available from the v10 OCO-2 MIP website.

For this study, the best estimate of NCE is taken to be the ensemble median for each experiment (denoted NCE\_{\text{experiment}}). The uncertainty in NCE is calculated as an estimate (denoted σ_{\text{NCE}}) of the distribution’s standard deviation using the interquartile range (IQR) of the v10 OCO-2 MIP ensemble. It is a robust estimate that requires only the middle 50% of the ensemble to be normally distributed (Hoaglin et al., 1985). Hence from the normal tables, to two decimal places,

\[
σ_{\text{NCE}} = \frac{\text{IQR}(\text{NCE})}{1.35}.
\]  

(4)

For country-level fluxes, the NCE estimates are first aggregated to country totals for each ensemble member before calculating the median and standard deviation. This is done because there are spatial covariances between 1° × 1° grid cells. Thus, first aggregating regions for each ensemble member accurately propagates the aggregate differences between regions across the ensemble members.

The NBE estimate is calculated by subtracting the ODIAC fossil fuel emissions from NCE. The variance in NBE is then taken to be the sum of the variances of NCE and FF:

\[
σ_{\text{NBE}}^2 = σ_{\text{NCE}}^2 + σ_{\text{FF}}^2.
\]  

(5)

3.2.1 Atmospheric CO₂ data included in v10 OCO-2 MIP

In situ CO₂ measurements (Fig. 3a and d) are drawn from five data collections made available in ObsPack format (Masarie et al., 2014). Those source ObsPacks and their references are listed in Table 2. These data include measurements from 55 international laboratories at 460 sites around the world. The majority of data are from the openly available GLOBALVIEW+ program but with some additional provisional data for 2020–2021 and data from other programs not participating in the GLOBALVIEW+ project. CO₂ measurements are broadly divided into two categories: those measurements we identify as suitable for assimilation and other measurements not suitable for assimilation.

In CO₂ inverse analyses, uncertainties ascribed to in situ measurements are a combination of the uncertainty in the measurement and a representativeness error from the inability of the forward model to accurately simulate the measurement (due to aspects like a coarse model grid). To characterize the representativeness error, we used an empirical scheme based on simulations from the v7 OCO-2 MIP (Crowell et al., 2019). In situ CO₂ measurements are simulated in a forward simulation, and then the model–data mismatch statistics are calculated to characterize the representativeness errors at each measurement location and for each season. Although this was the standard method for characterizing uncertainties for modeled in situ measurements, each v10 OCO-2 MIP group was free to choose how to set the uncertainties in their specific setups.

Of the in situ measurements designated as being appropriate for assimilation, about 5% were withheld for cross-validation purposes. These data were chosen to be as independent as possible from the measurements that were assimilated. For quasi-continuous measurements, such as those taken every 15 min at NOAA tall towers, measurements were withheld for entire days: we chose 5% of the days in the dataset, and we withheld every assimilable measurement on that day. This is also how CO₂ measurements on National Institute for Environmental Studies (NIES) ships were treated. Entire aircraft profiles in the NOAA light-aircraft profiling network are assumed to consist of vertically correlated measurements, so entire profiles were withheld: we chose 5% of aircraft profiles to withhold. Most flask sites have measurement sampling protocols intended to ensure independence;
they are often taken at weekly or biweekly intervals during meteorological conditions meant to allow regional background air masses to be sampled. Thus, we chose to withhold 5% of assimilable flask measurements. We also verified that datasets at the same site were withheld on the same days; aircraft profiles over tower sites were, for instance, withheld on the same days that tower data were withheld.

OCO-2 land (Fig. 3b and e) and ocean (Fig. 3c and f) $X_{CO_2}$ retrievals are performed using version 10 of NASA's ACOS full-physics retrieval algorithm (O’Dell et al., 2018). A common set of OCO-2 retrieval “super-obs” data were derived from these retrievals and were assimilated by each modeling group. These super-obs are obtained by aggregating retrievals into 10 s averages (which better match the coarse transport models’ grid cells used in the inversions) following the same procedure as the v9 OCO-2 MIP (Peiro et al., 2022). Specifically, individual scenes within the 10 s span are weighted according to the inverse of the square of the $X_{CO_2}$ uncertainty (standard deviations) produced by the retrieval, and correlations of +0.3 for land scenes and +0.6 for ocean scenes are assumed when calculating the uncertainty on the 10 s averages (see Sect. 3.2.1 of Baker et al., 2022); transport model errors are also considered (based on Schuh et al., 2019). Only 10 s spans with 10 or more good quality retrievals were used (sparser data being thought to be more prone to cloud-related biases). In the same vein as was done for the in situ data, $X_{CO_2}$ data from 5% of the orbits (entire orbits were withheld), chosen at random, were withheld for evaluation purposes.

### 3.3 Lateral carbon fluxes

Lateral carbon flux datasets (Table 3) include country-level $F_{river\ export}$ (Sect. 3.3.1), country-level $F_{crop\ trade}$ and country-level $F_{wood\ trade}$ (Sect. 3.3.2). Gridded lateral fluxes are estimated using a somewhat different approach and are described in Sect. 3.3.3.
3.3.1 Country-level $F_{\text{rivers export}}$

Rivers transport carbon laterally across land regions (e.g., to a lake) and from the land to the ocean. This lateral transport must be accounted for to quantify the total change in terrestrial carbon in a given region. However, there is considerable uncertainty in lateral carbon flux by rivers. To account for this, we use two independent estimates of country-level totals: one from the Dynamic Land Ecosystem Model (DLEM; Tian et al., 2010, 2015a) and the other based on Deng et al. (2022), who use the Global NEWS model (Mayorga et al., 2010) and observations across COastal Segmentation and related CATchments (COSCATs; Meybeck et al., 2006) that include dissolved inorganic carbon (DIC) of atmospheric origin, dissolved organic carbon (DOC) and particulate organic carbon (POC). These datasets cover 2015–2019. For 2020, we impose the 2015–2019 mean.

The DLEM is a process-based terrestrial ecosystem model that couples biophysical, soil biogeochemical, plant physiological and riverine processes with vegetation and land-use dynamics to simulate and predict the vertical fluxes, lateral fluxes, and storage of water, carbon, GHGs, and nutrient dynamics in terrestrial ecosystems and their interfaces with the atmosphere and land–ocean continuum (Tian et al., 2010, 2015a). There are three major processes involved in simulating the export of water, carbon and nutrients from land surface to the coastal ocean: (1) the generation of runoff and leachates; (2) the leaching of water, carbon, and nutrients from land to river networks in the form of overland flow and base flow; and (3) transport of riverine materials along river channels from upstream areas to coastal regions. The key processes and parameterization in the DLEM have been described in previous publications regarding the water discharge (Liu et al., 2013; Tao et al., 2014), riverine carbon fluxes (Ren et al., 2015, 2016; Tian et al., 2015b; Yao et al., 2021) and riverine nitrogen fluxes (Yang et al., 2015; Tian et al., 2020) from the terrestrial ecosystem to coastal oceans. The newly improved DLEM aquatic module better addresses processes within global small streams, which were recognized as hotspots of GHG emissions (Yao et al., 2020, 2021). DLEM produces estimates of the land loadings of carbon species (DIC, DOC and POC), CO$_2$ degassing and carbon burial during transporting, and the exports of carbon (DIC, DOC and POC) to the ocean for 105 basin-level segmentations (modified from COSCATs) (Meybeck et al., 2006). To estimate country totals, we map the basin carbon loss across land by assuming that the net carbon flux occurs uniformly across each basin. We then use the country mask to estimate the country totals for each region.

Deng et al. (2022) estimate the lateral carbon export by rivers to the coast minus the imports from rivers entering in each country (for relevant cases), including DOC, POC and DIC of atmospheric origin. Estimates of DOC, POC and DIC are obtained from the Global NEWS model (Mayorga et al., 2010), with a correction based on Resplandy et al. (2018) so that the global total exported to the coastal ocean is 2.86 Pg CO$_2$ yr$^{-1}$ (0.78 Pg C yr$^{-1}$). Deng et al. (2022) perform a correction to the Global NEWS estimates to remove the contribution of lithogenic carbon, using the methodology of Ciais et al. (2021).

For the analysis that follows, we estimate country-level totals of riverine lateral carbon fluxes by combining the estimates of DLEM with those of Deng et al. (2022). We take the mean of the two estimates to be the best estimate and take the magnitude of the difference between the estimates to be the 1σ uncertainty. Figure S1 shows the 2015–2019 mean annual net riverine lateral carbon fluxes. Fluxes are uniformly negative, implying a net flux of carbon from the land to the ocean and reduction in stored carbon for all countries. Fluxes are most negative in tropical rain forest and tropical monsoon climates, and they are smallest in more arid regions.

3.3.2 Country-level $F_{\text{wood trade}}$ and $F_{\text{crop trade}}$

Wood and crop products are traded between nations. We estimate the annual lateral fluxes of carbon due to this trade following the approaches of Deng et al. (2022) and Ciais et al. (2021). This approach utilizes crop and wood trade data compiled by the Food and Agriculture Organization of the United Nations (FAO; http://www.fao.org/faostat/en/#data, last access: 6 February 2023). The crop flux was estimated from the annual trade balance of 171 crop commodities calculated for each country. For wood products, we use the bookkeeping model of Mason Earles et al. (2012) to calculate the fraction of imported carbon in wood products that is oxidized in each country (for relevant cases), including DOC, POC and DIC of atmospheric origin. Estimates of DOC, POC and DIC are obtained from the Global NEWS model (Mayorga et al., 2010), with a correction based on Resplandy et al. (2018) so that the global total exported to the coastal ocean is 2.86 Pg CO$_2$ yr$^{-1}$ (0.78 Pg C yr$^{-1}$). Deng et al. (2022) perform a correction to the Global NEWS estimates to remove the contribution of lithogenic carbon, using the methodology of Ciais et al. (2021).

For the analysis that follows, we estimate country-level totals of riverine lateral carbon fluxes by combining the estimates of DLEM with those of Deng et al. (2022). We take the mean of the two estimates to be the best estimate and take the magnitude of the difference between the estimates to be the 1σ uncertainty. Figure S1 shows the 2015–2019 mean annual net riverine lateral carbon fluxes. Fluxes are uniformly negative, implying a net flux of carbon from the land to the ocean and reduction in stored carbon for all countries. Fluxes are most negative in tropical rain forest and tropical monsoon climates, and they are smallest in more arid regions.

### Table 3. Data sources for lateral flux estimates.

<table>
<thead>
<tr>
<th>Resolution</th>
<th>Flux</th>
<th>Model/data source</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>National</td>
<td>$F_{\text{rivers export}}$</td>
<td>Dynamic Land Ecosystem Model (DLEM) and Global NEWS with COSCATs data</td>
<td>Sect. 3.3.1</td>
</tr>
<tr>
<td>National</td>
<td>$F_{\text{wood trade}}$</td>
<td>UN FAO</td>
<td>Sect. 3.3.2</td>
</tr>
<tr>
<td>National</td>
<td>$F_{\text{crop trade}}$</td>
<td>UN FAO</td>
<td>Sect. 3.3.2</td>
</tr>
<tr>
<td>1° × 1°</td>
<td>$F_{\text{rivers export}}$</td>
<td>Global NEWS with COSCATs data</td>
<td>Sect. 3.3.3</td>
</tr>
<tr>
<td>1° × 1°</td>
<td>$F_{\text{wood trade}}$</td>
<td>UN FAO with downscaling</td>
<td>Sect. 3.3.3</td>
</tr>
<tr>
<td>1° × 1°</td>
<td>$F_{\text{crop trade}}$</td>
<td>UN FAO with downscaling</td>
<td>Sect. 3.3.3</td>
</tr>
</tbody>
</table>
assume fluxes equal to the 2015–2019 mean. The net crop and wood lateral fluxes and their uncertainties are shown in Fig. S2.

3.3.3 $1^\circ \times 1^\circ$ lateral flux estimates

Lateral fluxes at a higher resolution ($1^\circ \times 1^\circ$) follow similar principles to national values but were estimated separately with different implementation choices. High-resolution proxy data (satellite-derived NPP, population or livestock maps, etc.) enabled sub-national disaggregation. This was done using national totals based on FAO statistics for $F_{\text{wood trade}}$ and $F_{\text{crop trade}}$. For $F_{\text{rivers export}}$ these estimates were generated from Global NEWS and COSCATs data (DLEM was only used for national totals). For each $1^\circ \times 1^\circ$ grid cell, we assume the standard deviation of the mean flux to be 30% for $F_{\text{wood trade}}$ and $F_{\text{crop trade}}$ and 60% for $F_{\text{rivers export}}$. These uncertainty estimates are based on expert opinion, as a rigorous error budget has not yet been developed for the $1^\circ \times 1^\circ$ lateral flux estimates.

3.4 Estimate of carbon stock loss ($\Delta C_{\text{loss}}$)

Finally, we calculate $\Delta C_{\text{loss}}$ using Eq. (2) with the datasets described above. Assuming that the components contributing to $\Delta C_{\text{loss}}$ are independent, we calculate the uncertainty on $\Delta C_{\text{loss}}$ by combining the uncertainties (1 standard deviations) from the component fluxes in quadrature:

$$\sigma_{\Delta C_{\text{loss}}}^2 = \sigma_{\text{NBE}}^2 + \sigma_{\text{crop trade}}^2 + \sigma_{\text{wood trade}}^2 + \sigma_{\text{rivers export}}^2. \quad (6)$$

4 Evaluation of v10 OCO-2 MIP experiments

The performance of top-down CO$_2$ flux estimates can be impacted by a number of factors, including biases in the assimilated data, model transport, prior constraints and inversion architectures. Therefore, evaluating the performance of v10 OCO-2 MIP fluxes against independent observational datasets is critical for assessing high-quality flux estimates. Here, we evaluate the v10 OCO-2 MIP experiments in two ways. First, we compare the posterior CO$_2$ fields against independent CO$_2$ measurements (Sect. 4.1). Second, we compare the inferred air–sea CO$_2$ flux against estimates based on surface ocean CO$_2$ partial pressure ($p$CO$_2$) measurements (Sect. 4.2).

4.1 Evaluation of posterior CO$_2$ fields

We consider four atmospheric CO$_2$ datasets:

1. Withheld in situ CO$_2$ measurements. These are measurements contained in the ObsPack collection described in Sect. 3.2.1 but intentionally withheld for evaluation purposes. Independence from the assimilated data is ensured following the steps described in Sect. 3.2.1.

2. X$_{\text{CO}_2}$ retrievals from the TCCON. These data are acquired from a network of ground-based Fourier transform spectrometers measuring direct solar spectra from which X$_{\text{CO}_2}$ is retrieved (Wunch et al., 2011). For this analysis, we include 30 TCCON sites listed in Table A1. These data are filtered and aggregated following the method outlined in Appendix C of Crowell et al. (2019).

3. Withheld OCO-2 land glint and land nadir X$_{\text{CO}_2}$ retrievals. These data could have been assimilated, but they are intentionally withheld for evaluation purposes (Sect. 3.2.1).

4. Withheld OCO-2 ocean glint X$_{\text{CO}_2}$ retrievals. These data could have been assimilated, but they are intentionally withheld for evaluation purposes (Sect. 3.2.1).

We first perform a simple check on the inversion results by comparing the atmospheric CO$_2$ growth rate estimated from the v10 OCO-2 MIP experiments to that derived directly from NOAA CO$_2$ measurements (Fig. 4). The growth rate is estimated from CO$_2$ measurements and model co-samples at “marine boundary layer” sites, which predominantly observe well-mixed marine boundary layer air representative of a large volume of the atmosphere. A smooth curve is then fit to these data to estimate the global growth rate (Thoning et al., 1989). This is the same method employed by NOAA to report the CO$_2$ growth rate (http://www.gml.noaa.gov/ccgg/trends/, last access: 6 February 2023). We estimate the uncertainty in the measurement-based growth rate from the difference between the growth rate estimated here and that reported on the NOAA website. Differences between these estimates are primarily driven by differences in measurement sampling used for the website relative to that used here (as we are limited to withheld co-samples here). We calculate the uncertainty as the standard error of the mean for the differences between the growth rates estimated here and by NOAA across 2015–2019. This gives an uncertainty on the 5-year growth rate of $\pm 0.053$ ppm yr$^{-1}$. Note that NOAA reports the growth rate using the X2019 scale, whereas our estimates here are from the X2007 scale, which may contribute to the differences. We find that the IS, LNLG and LNLGIS experiments show good agreement with the NOAA estimate over this period. However, both the OG and LNLLOGIS experiments are found to have a high bias. This suggests that there may be a spurious trend in the v10 OCO-2 ocean glint X$_{\text{CO}_2}$ retrievals of 0.04–0.13 ppm yr$^{-1}$ (OG experiment bias) that impacts flux estimates in both experiments that assimilate ocean glint data.

Second, we estimate the overall observation–model agreement as the root mean square error (RMSE) for the withheld in situ CO$_2$, TCCON X$_{\text{CO}_2}$, withheld OCO-2 land X$_{\text{CO}_2}$ and withheld OCO-2 ocean X$_{\text{CO}_2}$ (Fig. 5). For the in situ and OCO-2 data, the normalized RMSE is shown, meaning that the observation–model difference is divided by the observational uncertainty (1σ). Overall, we find reasonably good agreement between the evaluation datasets and poste...
The estimates of the CO$_2$ growth rate for each experiment are computed by sampling the model CO$_2$ fields at the same times and locations as those used to derive the NOAA measurement-based estimate. Each v10 OCO-2 MIP experiment is shown as a box plot, with the error bars showing the full range, the shaded region showing the interquartile range and the solid line showing the median ensemble member of the ensemble.

Finally, we examine the mean bias over 2015–2020 for 30$^\circ$ latitude bins (Fig. 6). Similar to previous comparisons, we find that the OG experiment stands out as being more biased against the independent observations relative to the other experiments. In particular, the observation–model difference for the OG experiment tends to be lower (higher modeled CO$_2$) than the evaluation datasets. This is particularly evident in the northern extratropics. Over 30–60$^\circ$ N, where independent observations are densest, we find that the OG ensemble median is biased by approximately $-0.69$ ppm against TCCON, $-0.74$ ppm against withheld in situ and $-0.48$ ppm against withheld OCO-2 LNLG, suggesting a possible meridional bias (higher retrieved X$_{CO_2}$ than independent observations) in the OCO-2 ocean X$_{CO_2}$ retrievals. The IS, LNLG and LNLGIS experiments tend to show similar observation–model differences, suggesting limited ability to distinguish between the performance of these inversions in large-scale features.

All experiments show some biases against TCCON sites. In particular, low biases (high modeled CO$_2$) are found for 0–30$^\circ$ S and 60–90$^\circ$ N. The underlying cause for these differences is unknown. Figure S3 shows the monthly mean observation–model differences for each TCCON site and each experiment. The differences can be quite variable between sites but are generally similar between experiments (for IS, LNLG and LNLGIS). Some of these differences may be related due to representativeness errors, particularly for urban sites. For example, Caltech and JPL are within Los Angeles County and show a large positive bias, while nearby Edwards is less impacted by urban emissions and shows a much smaller bias (Schuh et al., 2021). However, other differences are harder to explain, such as a negative trend in the observation–model bias for Park Falls and positive at Darwin during the 2015–2020 period. Site-to-site biases among TCCON sites may also contribute to these differences.

Overall, this analysis finds that the OG experiment shows the poorest agreement against the evaluation datasets (excluding the withheld ocean glint data). The LNLGOGIS experiment shows the second worst performance against evaluation datasets, while the remaining experiments (IS, LNLG and LNLGIS) all show good agreement against the evaluation data. These results suggest that there may be residual biases in the OCO-2 ocean glint dataset that adversely impact the OG and LNLGOGIS experiments.
4.2 Comparison of air–sea fluxes with $p$CO$_2$-based estimates

The exchange of CO$_2$ between the atmosphere and the ocean (air–sea flux) can be estimated from measurements of the surface ocean partial pressure of CO$_2$ ($p$CO$_2$). These $p$CO$_2$ data are extrapolated to global maps and combined with gas transfer velocity parameterizations to infer global maps of the air–sea CO$_2$ fluxes (Fay et al., 2021). Although significant uncertainties remain, particularly in accurately representing the gas transfer velocity (Fay et al., 2021), comparisons between the $p$CO$_2$-based air–sea fluxes and v10 OCO-2 MIP experiments can inform possible biases between estimates and inform potential areas for future research.

Here, we compare v10 OCO-2 MIP air–sea fluxes to an ensemble of air–sea flux estimates from SeaFlux (Fay et al., 2021; Gregor and Fay, 2021a). SeaFlux developed a standardized approach to harmonize and extend six air–sea CO$_2$ flux products from as many surface $p$CO$_2$ products: JENA-MLS (Rödenbeck et al., 2013), MPI-SOMFFN (Land-schützer et al., 2014, 2020), CMEMS-FFN (Denvil-Sommer et al., 2019; Chau et al., 2022), CSIR-ML6 (Gregor et al., 2019), JMA-MLR (Iida et al., 2021) and NIES-FNN (Zeng et al., 2014). For each $p$CO$_2$ product, we examine the mean of six air–sea fluxes obtained using different wind reanalysis datasets to estimate the gas transfer parameterization (ERA5, JRA-55 and CCMP2). The spread among these six estimates provides a measure of uncertainty in the extrapolation of $p$CO$_2$ data to a global grid but does not account for errors in the gas transfer velocity formulation nor the uncertainties in the reanalysis winds used as input (Fay et al., 2021). Note that the prior estimates of air–sea CO$_2$ fluxes in v10 OCO-2 MIP experiments are generally $p$CO$_2$-based flux estimates and therefore not independent from the SeaFlux datasets.

Figure 7 shows the 2015–2019 mean air–sea fluxes for each of the six SeaFlux products and for the v10 OCO-2 MIP experiments across 30$^\circ$ latitude bands and large ocean regions. Over the global ocean, the $p$CO$_2$-based air–sea fluxes tend to give stronger removals (median $= -10.0$ Pg CO$_2$ yr$^{-1}$ or $-2.7$Pg C yr$^{-1}$, range $= -9.2$ to $-12.9$Pg CO$_2$ yr$^{-1}$ or $-3.5$ to $-2.5$Pg C yr$^{-1}$) than the v10 OCO-2 MIP, which range from $-7.9 \pm 1.9$Pg CO$_2$ yr$^{-1}$ ($-2.1 \pm 0.5$Pg C yr$^{-1}$) for the IS experiment to $-10.2 \pm 1.28$Pg CO$_2$ yr$^{-1}$ ($-2.8 \pm 0.4$Pg C yr$^{-1}$) for the OG experiment. On regional scales, the v10 OCO-2 MIP experiments overlap with the $p$CO$_2$-based estimates except for the northern high latitudes (60–90$^\circ$ N), where $p$CO$_2$-based estimates suggest systematically larger removals. Similarly, the $p$CO$_2$-based estimates tend to give greater removals over the southern midlatitudes (20–50$^\circ$ S).

The different v10 OCO-2 MIP experiments tend to give similar air–sea fluxes, except for the OG experiment in the tropics. Although not systematic, the OG experiment suggests weaker emissions in the tropics of 0.2±1.3 Pg CO$_2$ yr$^{-1}$ (0.05±0.34 Pg C yr$^{-1}$) relative to the median $p$CO$_2$-based estimate of 1.6 Pg CO$_2$ yr$^{-1}$ (0.43 Pg C yr$^{-1}$) with a range of 0.4 to 1.8 Pg CO$_2$ yr$^{-1}$ (0.10 to 0.50 Pg C yr$^{-1}$). Thus, similar to the evaluation of posterior CO$_2$ fields, the OG experiment is an outlier among the v10 OCO-2 MIP experiments, fur-
ther supporting the possibility that residual biases may exist in the ocean glint $X_{CO_2}$ retrievals.

5 Metrics for interpreting country flux estimates

To aid users in interpreting top-down country-level flux estimates, we provide two metrics. The first metric is called the $Z$ statistic and quantifies the statistical agreement between the IS and LNLG NCE estimates and thus gives an indication of how robust flux estimates are across the v10 OCO-2 MIP experiments (Sect. 5.1). The second metric is called the fractional uncertainty reduction (FUR) and informs the impact of the assimilated CO$_2$ data on the estimated fluxes (Sect. 5.2).

5.1 $Z$ statistic

The $Z$ statistic is defined as

$$Z_{statistic} = \frac{NCE_{LNLG} - NCE_{IS}}{SD(NCE_{LNLG} - NCE_{IS})},$$

where the denominator represents the standard deviation of $NCE_{LNLG} - NCE_{IS}$ across the ensemble members. Differences in NCE and $\Delta C_{loss}$ between v10 OCO-2 MIP experiments can be considerable. As an example, Fig. 8a shows that differences between $NCE_{LNLG}$ and $NCE_{IS}$ are notable for South America and Africa. The LNLG experiment gives more positive $\Delta C_{loss}$ (carbon loss from land) over northern sub-Saharan Africa and northeast South America but more negative $\Delta C_{loss}$ over southern tropical Africa, southern and eastern South America, and southeast Asia. We examine the $Z$ statistic (Fig. 8b) to quantify the statistical significance of these differences (magnitude greater than 1.96 indicates statistically significant differences at level $\alpha = 0.05$). Most countries do not have statistically significant differences, indicating relatively good agreement between the IS and LNLG ensembles. Significant differences primarily occur in small to mid-sized tropical countries. Canada also shows a systematic difference driven by small uncertainties in the IS and LNLG estimates.

5.2 Fractional uncertainty reduction (FUR)

Byrne et al. (2022) report the uncertainty in NCE as the standard deviation across v10 OCO-2 MIP ensemble members (estimated using Eq. 4). This metric incorporates uncertainties related to model transport and aspects of the inversion configuration, such as optimization technique and a priori flux estimates. However, this metric is different to the uncertainty metric usually computed in a Bayesian framework, that is, the Bayesian posterior uncertainty. That uncertainty quantifies the impact of errors in the observations and prior constraints on the posterior flux estimates. The Bayesian posterior uncertainty is not reported for practical reasons, as the majority of contributing models do not calculate this quantity, so it is not possible to calculate this quantity across the ensemble.
In this section, we examine the posterior uncertainty estimates from two contributing inversion systems (CAMS and TM5-4DVar) and compare these estimates to the ensemble-based uncertainty estimate provided with the dataset. Then, we define the FUR metric between the posterior and prior NCE estimates based on the TM5-4DVar model (as CAMS does not estimate uncertainties for the LNLGIS and LNLGOGIS experiments), which can be used to understand the relative impact of assimilated atmospheric CO$_2$ data on estimates of country-level NCE and $\Delta C_{\text{loss}}$.

Both CAMS and TM5-4DVar estimate CO$_2$ fluxes using four-dimensional variational assimilation (4D-Var) and estimate posterior uncertainty estimates using a Monte Carlo method derived by Chevallier et al. (2007). The realism of the prior and posterior CAMS uncertainty estimates has already been the topic of several studies (see Chevallier, 2021, and references therein). Figure 9 shows the ensemble-based uncertainty, prior/posterior uncertainty from CAMS (prior, IS and LNLG only) and prior/posterior uncertainty from TM5-4DVar for four countries. Notably, the magnitudes of the prior/posterior uncertainties from CAMS and TM5-4DVar are quite different, with CAMS uncertainties being 2–8 times larger. Differences in prior/posterior uncertainties of this magnitude are not unusual among inversion systems and highlight the sensitivity of Bayesian uncertainty estimates to choices about prior uncertainties. Both CAMS and TM5-4DVar posterior uncertainties are smaller relative to their prior by similar fractions, driven by the assimilated CO$_2$ data. The magnitude of the ensemble-based uncertainty tends to fall in between the CAMS and TM5-4DVar estimates. However, the CAMS and TM5-4DVar posterior uncertainty estimates decrease as more data are assimilated (as expected), while the ensemble spread does not. In fact, the ensemble spread increases with data density in some cases (e.g., Australia LNLGIS). Thus, overall, we find that the ensemble-based uncertainty estimate is of similar magnitude to the prior/posterior estimate but that the magnitude of posterior uncertainty is quite dependent on the assumed prior uncertainty.

We now calculate the FUR metric in NCE from the TM5-4DVar Bayesian uncertainties (note that we use TM5-4DVar only because CAMS does not report LNLGIS or LNLGOGIS uncertainties). FUR is calculated from the prior flux standard deviation ($\sigma_{\text{prior}}$) and posterior flux standard deviation ($\sigma_{\text{posterior}}$) as

$$FUR = 1 - \frac{\sigma_{\text{posterior}}}{\sigma_{\text{prior}}}.$$  

(8)

This quantity ranges between 0 and 1, with larger values indicating that the Bayesian uncertainties have decreased more (relative to the prior) due to the observational constraints from assimilated data. This metric is useful for understanding how the assimilation of data influences the NCE and $\Delta C_{\text{loss}}$ estimates, which may not be captured by the ensemble spread. For example, Saudi Arabia has a small NCE uncertainty estimate, but this is largely driven by prior knowledge that biosphere CO$_2$ fluxes, while the atmospheric CO$_2$ data have little impact on the NCE estimate.

Figure 10 shows FUR for the IS, LNLG, LNLGIS and LNLGOGIS experiments. FUR is larger in regions with denser observational coverage. For example, the IS FUR is close to 1 in the USA and parts of Europe, reflecting dense CO$_2$ measurements, but it remains small for many tropical countries, where sampling is sparse. Meanwhile, the LNLG experiment generally has larger FUR values than the IS experiment in the tropics, reflecting denser sampling, but has lower values for some small high-latitude countries, such as in Scandinavia.

6 Dataset description

The dataset described in this paper, Byrne et al. (2022), provides annual totals of country-level and $1^\circ \times 1^\circ$ gridded $\Delta C_{\text{loss}}$, NBE, NCE, $F_{\text{rivers export}}$ and the combined...
$F_{\text{crop trade}} + F_{\text{wood trade}}$ fluxes, as well as their uncertainties over 2015–2020. In addition, the country-level $Z$ statistic (Eq. 7) and FUR (Eq. 8) metrics are provided to help interpret the flux and stock change estimates. These data are provided for the v10 OCO-2 MIP IS, LNLG, LNLGIS and LNLGOGIS experiments. The OG experiment is excluded due to poor evaluation against independent CO$_2$ measurements and pCO$_2$-based air–sea fluxes, likely due to residual XCO$_2$ biases in the OCO-2 ocean glint XCO$_2$ retrievals (Sect. 4). We note that biases in ocean glint XCO$_2$ retrievals will also adversely impact flux estimates from the LNLGOGIS and caution against using these data when they show differences from the IS, LNLG and LNLGIS experiments. Future improvements to the OCO-2 XCO$_2$ retrievals are expected to reduce residual XCO$_2$ biases, and thus the quality of the LNLGOGIS experiment is expected to improve in future OCO-2 MIP experiments.

For the $1^\circ \times 1^\circ$ gridded dataset, we emphasize that caution is needed in interpreting these data. As discussed in Sect. 1.3, atmospheric CO$_2$ inversion analyses provide the
best constraints on the largest spatial scales (e.g., continental-to-global). The confidence in these top-down estimates decreases at smaller spatial scales. The minimum spatial resolution for robust flux estimates is dependent on the density and precision of the measurements and is challenging to quantify. However, scales smaller than France or Germany in geographic extent are unlikely to be meaningfully constrained. Thus, we recommend only using 1° × 1° CO2 fluxes aggregated to larger spatial scales. In aggregating, we recommend propagating uncertainties by assuming first 100 % correlation (sum of the 1° × 1° uncertainties) and then 0 % correlation (square root of the sum of the squared uncertainties) between grid cells. We strongly encourage contacting the authors before using the gridded 1° × 1° dataset.

These data are available for download from the Committee on Earth Observation Satellites’ (CEOS) website: https://doi.org/10.48588/npf6-sw92 (Byrne et al., 2022). The country-level data are available for download as comma-separated values (CSV), Network Common Data Form (NetCDF) and Microsoft Excel worksheet files. The 1° × 1° gridded dataset is available as a NetCDF file.

7 Characteristics of the dataset

Globally, over 2015–2020, we report FF emissions of 35.79 ± 1.50 Pg CO2 yr−1 (9.76 ± 0.41 Pg C yr−1), $F_{\text{rivers export}}$ of −3.35 ± 0.59 Pg CO2 yr−1 (−0.91 ± 0.16 Pg C yr−1), and globally balanced $F_{\text{crop trade}}$ and $F_{\text{wood trade}}$. Table 4 gives the global annual mean changes in the atmospheric burden of CO2, $\Delta C_{\text{gain}}$ and ocean sequestration. Across the experiments, the median fraction of fossil fuel emissions remaining in the atmosphere is 55 %–56 %, while 32 %–36 % is sequestered by the ocean and 9 %–13 % is sequestered by terrestrial ecosystems. Note that this omits land-use change (LUC) emissions of ∼3.85 Pg CO2 yr−1 (~1.05 Pg C yr−1). Friedlingstein et al., 2022), which are compensated for by additional carbon uptake by land. Of the combined FF+LUC emissions, 50 % remains in the atmosphere, 29 %–33 % is sequestered by the ocean and 18 %–21 % is sequestered by terrestrial ecosystems. Relative to the Global Carbon Budget 2021 (GCB 2021; Friedlingstein et al., 2022) we find 2.24–3.53 Pg CO2 yr−1 (0.61–0.96 Pg C yr−1) less removal by land (mean/median difference) but greater removal by the ocean of 0.87–2.24 Pg CO2 yr−1 (0.24–0.61 Pg C yr−1); however, these differences are consistent within 1 standard deviation of the mean/median values. Interestingly, we report greater removals by the ocean than GCB 2021 but reduced air–sea flux relative to SeaFlux. This can be explained by the fact that $pCO2$-based air–sea flux estimates generally give larger mean ocean carbon uptake than model estimates (Fay and McKinley, 2021) and that we estimate a larger $F_{\text{rivers export}}$ than GCB 2021.

Meridionally, NCE is largest in the northern extratropics, coinciding with the largest FF emissions (Fig. 11). However, the northern extratropics also show negative $\Delta C_{\text{loss}}$, implying increasing terrestrial carbon stocks, particularly between 30–60° N. NCE is less positive in the tropics, primarily due to lower FF emissions. However, this region tends to show neutral-to-positive $\Delta C_{\text{loss}}$, suggesting that terrestrial carbon stocks may be decreasing. The LNLG and IS results also differ most in the tropics, with LNLG suggesting greater terrestrial carbon stock loss over 0–30° N but less over 0–30°S. The differences in CO2 fluxes between these experiments are not well understood, and both experiments evaluate well against independent observations (Sect. 4).

The spatial distribution of NCE over 2015–2020 at 1° × 1° and aggregated to country scale for the LNLGIS experiment is shown in Fig. 12. At 1° × 1° (Fig. 12a and b), localized fossil fuel emissions are visible, generally corresponding to urban areas and industrialized regions. These emissions are interspersed over broad source and sink structures that are driven by biosphere removals or emissions. Land biosphere removal is most evident across the northern mid-high latitudes. In contrast, tropical removals and emissions are more regional. When NCE is aggregated to the country scale (Fig. 12c and d), most countries are net sources driven by fossil fuel emissions, particularly in the northern extratropics. Figure 12e–f show the 2015–2020 extent of these differences are likely impacted by the different spatial and temporal distribution of LNLG and IS measurements (see Sect. 5.2), model transport errors (Stephens et al., 2007; Schuh et al., 2019, 2022), and residual retrieval biases in the OCO-2 XCO2 retrievals (Peiro et al., 2022). Unfortunately, the regions showing the largest differences in fluxes generally have few independent atmospheric CO2 measurements for validation, limiting our ability to distinguish between different causes. Thus, we believe that NCE and $\Delta C_{\text{loss}}$ estimates are most reliable when agreement is found across the v10 OCO-2 MIP experiments.

Differences in NCE and $\Delta C_{\text{loss}}$ between the v10 OCO-2 MIP experiments can be considerable (the statistical significance of these differences is quantified by the Z statistic; see Sect. 5.1). The underlying cause of the differences between the v10 OCO-2 MIP experiments is not well understood, but the differences are likely impacted by the different spatial and temporal distribution of LNLG and IS measurements (see Sect. 5.2), model transport errors (Stephens et al., 2007; Schuh et al., 2019, 2022), and residual retrieval biases in the OCO-2 XCO2 retrievals (Peiro et al., 2022). Unfortunately, the regions showing the largest differences in fluxes generally have few independent atmospheric CO2 measurements for validation, limiting our ability to distinguish between different causes. Thus, we believe that NCE and $\Delta C_{\text{loss}}$ estimates are most reliable when agreement is found across the v10 OCO-2 MIP experiments.

We will now show examples of carbon budgets for four countries from this dataset. Figure 13 shows the 2015–2020 mean FF, $F_{\text{rivers export}}$, $F_{\text{crop trade}}$, $F_{\text{wood trade}}$, $\Delta C_{\text{loss}}$ and
Table 4. 2015–2020 global mean atmospheric increase, terrestrial carbon gain ($\Delta C_{\text{gain}}$) and ocean carbon gain from the IS, LNLG, LNLGIS and LNLLOGIS experiments (mean/median ± 1 standard deviation). Positive values of $\Delta C_{\text{gain}}$ and ocean carbon gain indicate increases in carbon stocks. GCB 2021 were obtained from the Global Carbon Budget 2021 (Friedlingstein et al., 2022) with $\Delta C_{\text{gain}}$ calculated as the difference between the land sink and land-use change emissions with errors propagated in quadrature.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Atmosphere</th>
<th>$\Delta C_{\text{gain}}$</th>
<th>Ocean carbon gain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(Pg CO$_2$ yr$^{-1}$)</td>
<td></td>
</tr>
<tr>
<td>IS</td>
<td>19.73 ±</td>
<td>4.58 ±</td>
<td>11.35 ±</td>
</tr>
<tr>
<td></td>
<td>0.19</td>
<td>2.44 ± 0.66</td>
<td>2.01 ± 0.55</td>
</tr>
<tr>
<td></td>
<td>(5.38 ± 0.05)</td>
<td>(3.12 ± 0.66)</td>
<td>(3.10 ± 0.55)</td>
</tr>
<tr>
<td>LNLG</td>
<td>19.64 ±</td>
<td>3.29 ±</td>
<td>12.91 ±</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
<td>2.93 ± 0.80</td>
<td>2.63 ± 0.72</td>
</tr>
<tr>
<td></td>
<td>(5.36 ± 0.02)</td>
<td>(9.09 ± 0.80)</td>
<td>(3.52 ± 0.72)</td>
</tr>
<tr>
<td>LNLGIS</td>
<td>19.64 ±</td>
<td>4.19 ±</td>
<td>11.98 ±</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>2.77 ± 0.75</td>
<td>2.32 ± 0.64</td>
</tr>
<tr>
<td></td>
<td>(5.36 ± 0.02)</td>
<td>(1.14 ± 0.75)</td>
<td>(3.27 ± 0.64)</td>
</tr>
<tr>
<td>LNLLOGIS</td>
<td>19.97 ±</td>
<td>4.03 ±</td>
<td>11.54 ±</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>2.36 ± 0.64</td>
<td>1.79 ± 0.49</td>
</tr>
<tr>
<td></td>
<td>(5.45 ± 0.05)</td>
<td>(1.10 ± 0.64)</td>
<td>(3.15 ± 0.49)</td>
</tr>
<tr>
<td>GCB 2021</td>
<td>19.8 ±</td>
<td>6.82 ±</td>
<td>10.67 ±</td>
</tr>
<tr>
<td></td>
<td>0.73</td>
<td>3.15 ± 0.86</td>
<td>1.83 ± 0.59</td>
</tr>
<tr>
<td></td>
<td>(5.39 ± 0.2)</td>
<td>(1.86 ± 0.86)</td>
<td>(2.91 ± 0.59)</td>
</tr>
</tbody>
</table>

Figure 11. Zonal mean (a) NCE, (b) FF + lateral fluxes and (c) $\Delta C_{\text{loss}}$ for 30° increments of latitude based on 1° x 1° estimates averaged over 2015–2020. IS, LNLG, LNLGIS and LNLLOGIS median estimates are shown by solid lines, and 1σ uncertainties are shown by the shaded region.

NCE fluxes for the USA, India, Indonesia and Australia. All of the CO$_2$ fluxes on the left of the dashed line combine to give the NCE flux constrained by the v10 OCO-2 MIP experiments. We find that FF is the strongest contributor to NCE for all countries but that $\Delta C_{\text{loss}}$ also plays a strong modulating role. For example, negative $\Delta C_{\text{loss}}$ (increasing terrestrial carbon stocks) for the USA reduces NCE to be less than would be expected given the FF emissions. Conversely, Indonesia has positive $\Delta C_{\text{loss}}$ (decreasing terrestrial carbon stocks), resulting in increased NCE relative to FF. Some countries also show differences in $\Delta C_{\text{loss}}$ between v10 OCO-2 MIP experiments. For example, the LNLG and LNLGIS experiments suggest negative $\Delta C_{\text{loss}}$ for India, while the IS suggests $\Delta C_{\text{loss}}$ is roughly neutral. Figures of carbon budgets for 28 additional countries (Fig. S8) and 14 regions (Fig. S9) are shown in the Supplement. The carbon budgets can also be examined for individual years (Fig. 14). Both Indonesia and Australia show considerable variations in $\Delta C_{\text{loss}}$ that drive variations in NCE over this period. Indonesia has a large positive $\Delta C_{\text{loss}}$ in 2015, driven by warm, dry weather and fires during the 2015 El Niño (Yin et al., 2016). Australia showed strong neg-
ative $\Delta C_{\text{loss}}$ (except for IS) during 2016, which was the 15th wettest year on record (precipitation 17% above average; Bureau Of Meteorology, 2017). Australia also showed anomalous positive $\Delta C_{\text{loss}}$ during 2019, which was the warmest and driest year on record, with considerable terrestrial carbon loss related to biomass burning in the southeast (Byrne et al., 2021). Variations in NCE are also found to be related to FF emissions. In particular, a reduction in NCE is found for 2019 and 2020 in the USA that is primarily linked to a reduction in FF emissions rather than $\Delta C_{\text{loss}}$. Time series of NCE and $\Delta C_{\text{loss}}$ for 28 additional countries (Figs. S10, S11) and 14 regions (Figs. S12, S13) are shown in the Supplement.

8 Comparison with national inventories

We demonstrate how the dataset presented here can be compared with NGHGIs reported under the UNFCCC, which were downloaded from https://di.unfccc.int/flex_annex1 (last access: 6 February 2023). We also refer the reader to
Figure 13. CO₂ budget for the USA, India, Indonesia and Australia averaged over 2015–2020. Bars show the median ± 1 standard deviation of FF, $F_{\text{rivers export}}$ ($R$), $F_{\text{crop trade}} + F_{\text{wood trade}}$ (CW), $\Delta C_{\text{loss}}$ and NCE (note that these quantities are related through Eq. 2).

Figure 14. Time series of the carbon budget for the USA, India, Indonesia and Australia. Solid lines show the median estimates, and shaded areas show ± 1 standard deviation.

Chap. 6.10.2 in vol. 1 of IPCC (2019) for additional discussion of comparing top-down estimates with NGHGIs. The fossil fuel emissions in Byrne et al. (2022) can be compared with the combined emissions from the energy and IPPU (Energy+iPPU) categories. In both cases, these estimates account for anthropogenic CO₂ emissions from the burning of fossil fuels and production of cement and other materials. We expect these estimates to generally be in good agreement, as they are similarly based on bottom-up accounting for national totals. However, the estimates may diverge when there are missing activity data, particularly in non-Annex 1 countries and more recent years (Andrew, 2020).

$\Delta C_{\text{loss}}$ can be compared to the combined emissions and removals from the agriculture, LULUCF and waste (Agr+LULUCF+Waste) categories. These quantities are not identical, with the most important difference being that NGHGIs are only for managed land, while $\Delta C_{\text{loss}}$ includes both managed and unmanaged lands. Therefore, caution is needed for parties with large unmanaged land areas (e.g., Canada or the Russian Federation). Another difference from
NGHGI is that $\Delta C_{\text{loss}}$ implicitly includes deposition of carbon in water body sediments within a country (such as lakes). However, this is expected to be a small contribution. Similarly, volcanic CO$_2$ emissions are implicitly included in $\Delta C_{\text{loss}}$ but are also believed to be small contributions (global subaerial volcanic CO$_2$ emissions are $\sim 0.05$ Pg CO$_2$ yr$^{-1}$, Fischer et al., 2019). It is worth noting that NGHGI require estimates of turnover times for wood products in producing countries, as these can have lifetimes of decades to centuries (see Appendix 3a.1 of Penman et al., 2003). No such estimate is needed for the top-down methods, as emissions from decaying wood products will be implicitly incorporated in NCE. Therefore, top-down methods only need to account for the lateral movement of wood products from the region where the carbon is sequestered to the region where the wood products are used and decompose.

For this analysis, we compare NGHGI and our dataset for three entities: the USA, European Union plus the United Kingdom (EU27+UK) and Australia. These were chosen for two reasons. First, NCE is better constrained by atmospheric CO$_2$ data over these relatively large regions. This is reflected in the FUR metric, which gives values of 0.76–0.91 for the USA (meaning a 76%–91% uncertainty reduction), 0.38–0.51 for EU27 and 0.45–0.78 for Australia. Second, each of these entities has small unmanaged land areas, making this more of an apples-to-apples comparison. An area of 95% of the USA is managed, with most unmanaged land being in the state of Alaska (Ogle et al., 2018). Similarly, all land in the EU27+UK is considered managed except for 5% of France’s territory (Petrescu et al., 2021).

Figure 15 shows time series of emissions and removals from NGHGI and Byrne et al. (2022) over 2015–2020. We focus our analysis on the 2015–2020 mean estimates, as emissions from decaying wood products are used and decompose. Averaged over the 2015–2020 period, we obtain statistically significant differences of 0.59–0.91 Pg CO$_2$ yr$^{-1}$ for the USA and 0.99–1.79 Pg CO$_2$ yr$^{-1}$ for the EU27+UK. The reasons for these differences are unclear but are not expected to be explained by removals in unmanaged lands. It is possible that NGHGI methods miss or underestimate sink processes and/or that there are biases affecting the top-down estimates (see Sect. 9 for remaining challenges in top-down estimates). We encourage further research and comparison between the NGHGI and top-down research communities to better understand the sources of these differences.

9 Discussion

Here we discuss the current limitations of top-down country-level CO$_2$ budgets and activities that can improve these estimates. Section 9.1 discusses current CO$_2$ observing systems and possible future expansions. Section 9.2 discusses current atmospheric CO$_2$ inversion systems, planned developments and opportunities for improvement. Finally, Sect. 9.3 discusses remaining challenges in estimating carbon stock changes from atmospheric CO$_2$ inversions.

9.1 Observations

In the context of global inversion analyses, annual mean biosphere–atmosphere CO$_2$ fluxes are best informed by measurements of atmospheric CO$_2$ on large spatial scales (e.g., continental-to-global) due to rapid mixing in the atmosphere and gaps in current measurement coverage. The confidence in these top-down estimates decreases as we move to smaller spatial scales, with the minimum spatial scale being dependent on the density, precision and sensitivity of the measurements. Future refinements in top-down CO$_2$ budgets will depend on increasing observational density (Sect. 9.1.1), improved validation (Sect. 9.1.2) and data harmonization (Sect. 9.1.3).

9.1.1 Expanding observations

An expanding network of CO$_2$ observing systems provides an opportunity to reduce uncertainties in top-down estimates of NCE. Across much of the globe, country-level estimates of NCE have been limited by the observational coverage of in situ CO$_2$ measurements and $X_{CO_2}$ retrievals. However, there are a number of planned expansions in observing systems that will help fill data gaps.

The first generation of space-based CO$_2$ systems currently in operation (GOSAT, GOSAT-2, OCO-2, OCO-3, TanSat) were designed primarily as proof-of-concept missions to demonstrate that space-based measurements could yield $X_{CO_2}$ retrievals with the precision and accuracy required to quantify emissions and removals of CO$_2$. Planned future missions will expand and improve upon current observing systems. MicroCarb, a France–UK mission, is expected to start operations in 2023 with an additional spectral band to better characterize the light path for the estimation of $X_{CO_2}$ (Bertaux et al., 2020). Japan’s GOSAT-GW mission (https://gosat-gw.nies.go.jp/en/, last access: 6 February 2023), which will be launched in early 2025, will also incorporate improved capabilities for CO$_2$ and CH$_4$. In 2025, the European Copernicus program will begin to deploy the first operational CO$_2$ and CH$_4$ monitoring constellation, CO2M (Pinty et al., 2017; Janssens-Maenhout et al., 2020). The CO2M constellation will eventually include up to three satellites, flying in formation to collect measurements at 2 by 2 km resolution over the entire globe at weekly intervals. In
addition, a follow-on to the Chinese TanSat mission is currently under development (Yang et al., 2018).

Most current and planned space-based CO₂ observing systems are passive, in that they rely on reflected sunlight to retrieve \(X_{\text{CO}_2}\). Active satellite missions, which use lidars for their light source, could provide coverage when reflected sunlight is not available or of insufficient intensity, such as at night and at high latitudes in the winter hemisphere when solar zenith angles are large. These systems also have the potential to better characterize systematic errors in current passive instruments by using pulse timing information to get a better estimate of path length and to filter out scattered light from clouds and aerosols (Abshire et al., 2010).

As space-based CO₂ observing systems expand, sub-orbital discrete air sampling (i.e., flask) and continuous CO₂ observing systems will remain critical for developing top-down CO₂ budgets. These in situ observations are the global standard for GHG measurements, because they can undergo direct calibration relative to the World Meteorological Organization (WMO) CO₂-in-air mole fraction scale, which is International System of Units (SI)-traceable (Hall et al., 2021). In contrast, open-path remote sensing measurements (both TCCON and satellite) can not be calibrated using standard gasses; they can only be compared to in situ vertical profile observations made relative to the WMO scale, with the differences used to adjust the remote sensing observations (e.g., Wunch et al., 2011). As such, in situ data are critical for linking remote sensing observations of CO₂ to the accepted trace gas scales. In situ data also provide complementary observational coverage to space-based observing systems (Byrne et al., 2017). Space-based measurements have broad spatial coverage but with seasonal variations driven by sunlight and have data gaps in persistently cloudy regions. In contrast, flask and in situ data can be deployed year-round and regardless of cloud cover. Additionally, in situ observations most typically represent the planetary boundary layer where flux signals in atmospheric CO₂ are larger than the signal as expressed in the column mean (Feng et al., 2019). Thus, these data play a critical role in improving carbon cycle constraints, especially in high-latitude and persistently cloudy regions (such as the tropics), and we encourage an expansion of these systems in these undersampled regions. Regular measurements of CO₂ using light aircraft above several sites in Amazonia exist (e.g., Gatti et al., 2021; Miller et al., 2021), but these measurement records, as well as a nascent aircraft program in Uganda, have so far been funded using short-term grants.

Measurements of stable isotope \(^{13}\text{C}/^{12}\text{C}\) and radiotope \(^{14}\text{C}/\text{C}\) ratios of carbon in CO₂ provide powerful tools for source attribution. Radiocarbon is absent from fossil fuels, making it ideal for distinguishing fossil versus biologic carbon fluxes, and inversions using measurements of CO₂ and \(^{14}\text{C}/\text{C}\) have been used to provide top-down constraints on national-scale fossil CO₂ emissions (Basu et al., 2020). Atmospheric \(^{13}\text{C}/^{12}\text{C}\) ratios provide insight into ecosystem stress and its relation to climate via constraint of ecosystem water use efficiency (photosynthesis relative to water loss by transpiration) and have been used in box models (Keeling
et al., 2017) and inversions (Peters et al., 2018). Atmospheric 
\( ^{13}\text{C}/^{12}\text{C} \) ratio data are generally available where discrete air 
samples are collected by various networks, but \( ^{14}\text{C}/\text{C} \) ratio 
data are more limited as they tend to require larger samples 
and measurement costs are greater. Other tracers closely re-
lated to \( \text{CO}_2 \) and measurement costs are greater. Other tracers closely re-
data are more limited as they tend to require larger samples 
and measurement costs are greater. Other tracers closely re-
lated to \( \text{CO}_2 \), such as \( \text{O}_2/\text{N}_2 \) (Keeling and Graven, 2021) 
and carbonyl sulfide (e.g., Hu et al., 2021; Remaud et al., 
2022) are also limited yet provide valuable information on 
global ocean/NBE and regional-scale photosynthesis/respira-
tion partitioning, respectively. Increasing the temporal and 
spatial density of these data, particularly across poorly sam-
ples regions, will allow for more diagnostic power of carbon 
cycle processes than is possible with \( \text{CO}_2 \) alone.

9.1.2 Data validation

Validation of \( X_{\text{CO}_2} \) retrievals is critical for ensuring that re-
terval biases do not strongly impact flux estimates. Current 
gaps in coverage of ground-based and airborne measures-
ments have limited our confidence in flux inferences from 
space-based data. For example, large \( \text{CO}_2 \) emissions over 
 northern sub-Saharan Africa are a robust feature of the inver-
sions that assimilate satellite \( X_{\text{CO}_2} \) retrievals (Palmer et al., 
2019), but there are few independent \( \text{CO}_2 \) measurements to 
confirm whether this inference is a real signal or an artifact 
of regional retrieval biases. Increased validation of space-
based observations will also provide critical support for im-
proved space-based inferences. Space-based measurements 
relax on validation against ground-based \( X_{\text{CO}_2} \) retrievals from 
the TCCON (Wunch et al., 2011) and the COllaborative Car-
bon Column Observing Network (COCCON; Frey et al., 
2019). In turn, these sites rely on in situ \( \text{CO}_2 \) measurements from 
aircraft profiles and AirCore (Karion et al., 2010) to tie 
their measurements to the WMO scale (Wunch et al., 2010; 
Messerschmidt et al., 2011). These data have been critical 
for validating and improving \( X_{\text{CO}_2} \) retrievals (Wunch et al., 
2017b; O’Dell et al., 2018; Kiel et al., 2019). Continued 
funding of these activities will be crucial for improving top-
down \( \text{CO}_2 \) flux estimates, and expansion of these observing 
systems into undersampled regions, such as the tropics and 
high latitudes, will also be important for identifying and ad-
ressing residual \( X_{\text{CO}_2} \) retrieval biases. In addition, efforts 
to cross-calibrate TCCON and COCCON sites will be help-
ful for minimizing site-to-site biases and identifying spuri-
ous drifts in \( X_{\text{CO}_2} \). We encourage future campaigns aimed at 
site-to-site comparisons similar to the FRM4GHG campaign 
that deployed total column \( \text{GHG} \) traveling standard instru-
ments at several TCCON sites as part of ESA’s FRM4GHG-2 
project (Sha et al., 2020).

9.1.3 Data harmonization

Further advancements in top-down flux estimates will be 
possible through combining the observational constraints 
from the constellation of space-based sensors and ground-
based instruments. Assimilating these data concurrently 
within inversion systems will increase our ability to re-
cover net fluxes over smaller regions. However, these instru-
m ents must be cross-calibrated against common standards to 
use these data together, as small intercalibration differences 
could potentially strongly impact flux estimates. We encour-
age support of these critical cross-calibration activities, as are 
outlined in Crisp et al. (2018).

9.2 Atmospheric \( \text{CO}_2 \) inversions

Atmospheric \( \text{CO}_2 \) inversion analyses are a critical tool for 
estimating surface fluxes from observations of atmospheric 
\( \text{CO}_2 \). Expanding observational coverage provides both op-
portunities and challenges for inversion systems. By address-
ing the current limitations of our inversion systems, we will 
be able to take full advantage of increasing observations to 
 improve country-level top-down estimates of NCE and 
\( \Delta C_{\text{loss}} \). Here we discuss ongoing and planned developments 
(Sect. 9.2.1), improving model transport (Sect. 9.2.2), miss-
ing processes and required assumptions (Sect. 9.2.3), and un-
certainty quantification (Sect. 9.2.4).

9.2.1 Ongoing and planned developments

To date, there are four operational or quasi-operational at-
mospheric \( \text{CO}_2 \) inversion systems: CarbonTracker (Jacobson 
et al., 2020), CAMS (Chevallier et al., 2005b), Jena Car-
boScope (Rödenbeck et al., 2018) and CMS-Flux (J. Liu 
et al., 2021) that are regularly updated on annual or quarterly 
timescales. These systems produce NBE and air–sea flux 
estimates from either in situ \( \text{CO}_2 \) measurements (Carbon-
Tracker, Jena CarboScope), OCO-2 \( X_{\text{CO}_2} \) retrievals (CMS-
Flux) or both (CAMS). Similarly, there are seven inversion 
models (including the aforementioned models) that update 
\( \text{CO}_2 \) flux estimates annually for the Global Carbon Bud-
get (Friedlingstein et al., 2022), including CAMS (Chevallier 
et al., 2005b), CarbonTracker Europe (CTE; Van der Laan-
Luijkx et al., 2017), Jena CarboScope (Rödenbeck et al., 
2018), UoE in situ (Feng et al., 2016), NISMON-CO2 (Niwa 
et al., 2017), MIROC4-.ACTM (Saeki and Patra, 2017; Chandra 
et al., 2022) and CMS-Flux (J. Liu et al., 2021).

The OCO-2 MIP activities have semiregularly performed 
ensemble inversion experiments (Crowell et al., 2019; Peiro 
et al., 2022). To date, OCO-2 MIP experiments have been 
linked to new versions of the ACOS retrieval algorithm, with 
major improvements to the quality of \( X_{\text{CO}_2} \) retrievals occur-
ing during each update. However, as the quality of retrievals 
have improved (particularly for ACOS v10 onwards), up-
dates to the ACOS retrieval algorithm are becoming less of a 
driver for new OCO-2 MIP experiments. In the future, OCO-
2 MIP activities could become more regular with annual up-
dates.

The first top-down \( \text{CO}_2 \) system for use in inventory 
development is CarbonWatch-NZ, under development in
New Zealand (https://niwa.co.nz/climate/research-projects/carbon-watch-nz, last access: 6 February 2023). This program includes expanded CO₂ measurement sites and the development of a regional atmospheric CO₂ inverse system to quantify the carbon budgets of New Zealand’s forest, grassland and urban environments. Initial results suggest stronger uptake by intact forests than estimated through bottom-up estimates (Steinkamp et al., 2017). This system may serve as an example for other nations through the Integrated Global GHG Information System (IG3IS) framework.

Beyond existing activities, there are a number of planned projects. The European Commission’s Copernicus program (https://www.copernicus.eu, last access: 6 February 2023) has a number of developments ongoing and planned, particularly in building anthropogenic CO₂ emission monitoring and verification support capacity (CO2MVS; Janssens-Maenhout et al., 2020), which is directly linked to the development and launch of the new CO2M mission and is expected to be operational from 2026 onwards. Further, there are a number of recently completed, ongoing and planned projects to develop and improve inversion systems to develop operational capacity. Examples include the recently completed CO₂ Human Emissions (CHE) project (https://www.che-project.eu/, last access: 6 February 2023) and the follow-up CoCO2 project (https://coco2-project.eu/, last access: 6 February 2023) that is ongoing, as well as the VERIFY project (https://verify.lsce.ipsl.fr/, last access: 6 February 2023). These projects are developing and refining inversion systems to estimate anthropogenic fossil fuel emissions, as well as emissions and removals from the agriculture, LULUCF and waste categories. Future planned projects include developing approaches to utilize co-emitted species and auxiliary observations (¹⁴C, solar-induced fluorescence, carbonyl sulfide and others) in order to isolate some of the CO₂ budget components and improve our understanding of the carbon cycle. For example, multiple data streams could be used together to optimize the dynamic global vegetation model parameters (e.g., Peylin et al., 2016).

In contrast to recent European efforts, there is no mandate for an operational top-down carbon-flux-attribution system in the USA. Nevertheless, efforts at NOAA centered around CarbonTracker (Jacobson et al., 2020) have been able to produce NBE estimates with relatively low latency, harnessing the agency’s substantial flask and in situ CO₂ network. In addition, NOAA has developed a higher-spatial-resolution North American regional inverse system, CarbonTracker-Lagrange (https://gml.noaa.gov/ccgg/carbontracker-lagrange/, last access: 6 February 2023; Hu et al., 2019). In anticipation of the launch of OCO in 2009, NASA started supporting research and development efforts needed to prototype an operational flux estimation system. In particular, the Carbon Monitoring System program (https://carbon.nasa.gov/, last access: 6 February 2023) has led to the development of both low-latency (2 months) atmospheric CO₂ reanalysis (Weir et al., 2021a) and approaches to combine top-down NCE estimates with other trace gas measurements (e.g., CO) and non-atmospheric carbon data (e.g., above-ground biomass) to provide improved understanding of carbon cycle processes (Liu et al., 2017; Byrne et al., 2020, 2021; Bloom et al., 2020). There is substantial technical capacity to build an operational system, but this requires a coordinated effort between federal agencies, academia and private interests.

In Canada, a prototype operational regional inverse modeling system, the Environment and Climate Change Canada (ECCC) National Carbon Flux Inversion System (ENCIS), is being developed to provide quantitative information on CO₂ (and CH₄) flux estimates over Canada from national to provincial scales, as well as to understand the carbon cycle in Canada such as CO₂ flux in boreal managed and unmanaged forests, wetland emissions of CH₄, and GHG emissions over a potentially thawing permafrost in response to the climate change. ENCIS is a regional inverse modeling system based on a Lagrangian approach and driven by metrology from the Global Environmental Multiscale (GEM) model (Girard et al., 2014) and is expected to have 1° × 1° spatial resolution.

Finally, there are ongoing internationally organized activities. Phase 2 of the Regional Carbon Cycle Assessment and Processes project (RECCAP-2), coordinated by the Global Carbon Project (https://www.globalcarbonproject.org/reccap/, last access: 6 February 2023), has aimed to characterize regional carbon budgets. This included investigating how different data sources – including atmospheric inversion analyses – can contribute to this goal (Bastos et al., 2022; Deng et al., 2022). In addition, the WMO has hosted workshops and symposiums with the GHG monitoring community to develop a framework for sustained, internationally coordinated global GHG monitoring (e.g., https://community.wmo.int/meetings/wmo-international-greenhouse-gas-monitoring-symposium, last access: 6 February 2023).

### 9.2.2 Improving CTM transport

Errors in the representation of atmospheric transport by CTMs have long been recognized as a major source of error in atmospheric CO₂ inversion analyses (Law et al., 1996; Law and Simmonds, 1996; Denning et al., 1995, 1999a, b; D. Baker et al., 2006; Stephens et al., 2007). Improvements to model transport will provide critical improvements to NCE and ΔC_loss estimates. Systematic errors in model transport limit our ability to relate surface fluxes and CO₂ observations and can lead to incorrect inferences of surface fluxes (Yu et al., 2018; Schuh et al., 2019; Stanevich et al., 2020). Improving model transport will require work in two areas: (1) improving model parameterizations of unresolved transport, particularly in coarse offline CTMs (like GEOS-Chem run at 4° × 5° in this ensemble) where the spatial and tem-
poral coarsening of meteorological fields can “average out” vertical transport that is resolved in the parent model (Yu et al., 2018; Staneevich et al., 2020); and (2) increasing spatial and temporal resolution in model simulations, which can better resolve atmospheric transport processes (Augustí-Panareda et al., 2019; Schuh et al., 2019). However, it should be noted that there are limitations to the improvements that can come from increased model resolution in the global inversion context due to underlying meteorological uncertainties (Liu et al., 2011; Polavarapu et al., 2016, 2018; McNor-
ton et al., 2020). Computational cost is also a significant challenge in inversion systems, because transport models usually scale poorly on supercomputers, for example because of the volume of meteorological data required as input.

As transport models are refined, it will be critical to periodically test their ability to represent large-scale atmospheric dynamics. This can be tested using long-lived trace gas species, including sulfur hexafluoride (Schuh et al., 2019), idealized age of air tracer (Krol et al., 2018) and beryllium-7 (Staneevich et al., 2020). Simulations of these trace species are critical in the context of inversion MIPs to gauge intermodel variability and average model bias (Schuh et al., 2019). Similarly, $^{222}$Rn is a useful short-lived gas species that enables modelers to evaluate the vertical mixing within the column (Remaud et al., 2018). In addition, model intercomparison studies have proven useful for diagnosing transport errors (e.g., Gaubert et al., 2019; Zhang et al., 2022), and we recommend further activities, such as within the Atmospheric Tracer Transport Model Intercomparison Project (TRANSCOM) framework.

9.2.3 Missing processes and required assumptions

The flux estimates provided here do not explicitly account for the atmospheric chemical production of atmospheric CO$_2$, which occurs from the oxidation of reduced carbon gasses. Instead, these fluxes are either prescribed as surface–atmosphere fluxes (e.g., for FF CO emissions) or neglected from the prior fluxes. This can cause inverse modeling systems to implicitly incorporate the atmospheric CO$_2$ source in optimized surface–atmosphere emissions and removals (i.e., air–sea fluxes and NBE), which can be far from the actual source of the reduced gas. For example, FF CO emissions are largely emitted in the northern extratropics but transported to the tropical troposphere. These incorrectly located emissions of CO$_2$ are large enough to impact top-down inversions (Enting and Mansbridge, 1991; Sunharalingam et al., 2005; Nassar et al., 2010; Wang et al., 2020). Future studies that aim to incorporate an atmospheric source of CO$_2$ would help correct for this current spatial bias (Ciais et al., 2022).

A critical assumption in the top-down CO$_2$ budgets estimated here has been that FF emissions are known and unbiased. Uncertainties in inventory-based FF emission estimates at global and country levels (e.g., Andres et al., 2014) are smaller than top-down NCE estimates; however, inventory-based emission estimates are prone to systematic biases due to the nature of the estimation approach (Guan et al., 2012; Oda et al., 2019), and FF uncertainties could bias the partitioning of NCE between FF and NBE (and propagate into $\Delta C_{\text{loss}}$) over countries with large emissions and lower reliability of statistical data collection systems, such as China. For example, Saeki and Patra (2017) show that an inferred increase in removals of CO$_2$ by the biosphere over China during 2001–2010 is likely to be an artifact imposed by an error in the trend of anthropogenic CO$_2$ emissions.

9.2.4 Uncertainty quantification

The uncertainty in NCE reported here is an estimate of the standard deviation of the v10 OCO-2 MIP ensemble members. This is meant to characterize uncertainties originating from the inversion configuration (such as the transport model, inversion method and prior constraints). However, there are also limitations to this method. First, there is only a small ensemble of 11 MIP ensemble members included in this analysis and an over-representation of inversions using two transport models: TM5 (3) and GEOS-Chem (5), which makes uncertainty quantification challenging. Future approaches that employ “borrowing strength” (Mears and et al., 2007; Cressie and Kang, 2016) could be employed to better characterize ensemble uncertainty. Second, the ensemble-based estimate does not capture some sources of uncertainty. In particular, Bayesian posterior uncertainties are not considered here (see Sect. 5.2), due to the fact that many of the inversion systems participating in the v10 OCO-2 MIP do not calculate this uncertainty. In addition, we find that the ensemble members that produce Bayesian uncertainties show large differences in magnitude. Thus, this is an area of future improvement for MIP activities, and we recommend more work into characterizing this error component in ensemble inversion experiments. We also note that using an analytic framework, posterior uncertainties and their sensitivities to prior information could be further examined, as has been done for methane (Worden et al., 2022).

9.3 Stock change estimates

Agriculture and LULUCF emissions and removals are generally quantified as terrestrial carbon stock changes in managed lands. A number of challenges remain in estimating this quantity from top-down methods. Firstly, lateral fluxes of carbon remain quite uncertain (and associated uncertainty estimates are themselves quite uncertain). The best constrained lateral fluxes are annual country-level $F_\text{wood trade}$ and $F_\text{wood trade}$, which are reported to the UN Food and Agriculture Organization. These fluxes are more uncertain on sub-national scales and sub-annual timescales. Meanwhile, $F_\text{flovers export}$ is best quantified on basin scales, where stream gauge measurements inform carbon fluxes. Improving sub-

https://doi.org/10.5194/essd-15-963-2023
national and sub-annual estimates of lateral fluxes would have several benefits: first, this would allow for better sub-national attribution, where regional fluxes could be better quantified. Second, this would allow for incorporating the atmospheric imprint of these carbon fluxes as prior information within atmospheric CO₂ inversion analyses, which may improve flux estimates on sub-national scales.

The GST and Paris Agreement do not consider emissions and removals from unmanaged lands. Separating managed lands from unmanaged lands remains a major challenge, given the smoothed large-scale CO₂ flux constraints provided by these top-down methods and the fact that both managed and unmanaged lands can experience considerable stock changes driven by interannual climate variations (e.g., El Niño) and in response to rising CO₂ and climate change. In addition, separating managed and unmanaged lands is further complicated by the fact that there is considerable ambiguity in the definitions of managed lands, which can also vary by country (Grassi et al., 2018; Chevallier, 2021). We recommend that each party provide a mask to unambiguously define the plots considered managed from year to year (Chevallier, 2021).

10 Data availability

Top-down CO₂ budgets are available from the Committee on Earth Observation Satellites’ (CEOS) website: https://doi.org/10.48588/mpf6-sw92 (Byrne et al., 2022). Gridded NBE and air–sea fluxes from the OCO-2 MIP are available at https://gml.noaa.gov/ccgg/OCO2_v10mip/ (Baker et al., 2023). Fossil fuel emissions prescribed in the inversions can be downloaded from https://doi.org/10.5281/zenodo.4776925 (Basu and Nassar, 2021). The ODIAC2020 emission data product can be downloaded from the Global Environmental Database hosted by the Center for Global Environmental Research at NIES (https://doi.org/10.17595/20170411.001, Oda and Maksyutov, 2015). SeaFlux pCO₂-based air–sea fluxes were downloaded from https://doi.org/10.5281/zenodo.5482547, (Gregor and Fay, 2021b).

11 Conclusions

We introduced a pilot top-down CO₂ budget dataset (Byrne et al., 2022) intended to start a dialogue between research communities and to identify ways that top-down flux estimates can inform country-level carbon budgets. This dataset provides annual country-level and 1° × 1° gridded top-down NCE and ΔCloss over 2015–2020, in addition to bottom-up FF and lateral fluxes. These data are provided for four experiments from the v10 OCO-2 MIP that differ in the data used in the assimilation: IS, LNLG, LNLGIS and LNLGOGIS. In addition, we provide two metrics for interpreting country-level estimates: (1) the Z statistic (Sect. 5.1), which quantifies the agreement between IS and LNLG NCE estimates, and (2) the FUR (Sect. 5.2), which quantifies the impact of atmospheric CO₂ data in reducing flux uncertainties.

Country-level flux estimates generally show robust signals for large extratropical countries (e.g., USA, Russia, China). Agreement between the experiments generally decreases for mid-sized countries (e.g., Turkey), particularly in regions with sparse observational coverage for the in situ network (such as the tropics). Large divergences between the IS and LNLG experiments occur in some regions, particularly in northern sub-Saharan Africa, and could be related to the sparsity of in situ CO₂ measurements or biases in OCO-2 retrievals. However, the sparsity of independent CO₂ measurements in these regions precludes definitive conclusions. We urge caution in interpreting the 1° × 1° gridded results and suggest collaborating with experts in atmospheric CO₂ inversion systems when using those data.

The accuracy of top-down NCE estimates was characterized through comparisons against independent atmospheric CO₂ datasets and through comparisons against pCO₂-based air–sea CO₂ fluxes. Overall, the IS, LNLG and LNLGIS were found to show the best agreement against independent CO₂ measurements, and we recommend using these experiments for analysis. Poorer agreement was found for experiments assimilating OCO-2 ocean glint XCO₂ retrievals, suggesting that residual retrieval biases adversely impact the LNLGOGIS experiment, and we urge caution in interpreting these data.

For future GSTs, top-down NCE estimates will be refined as new space-based XCO₂ observing systems expand and retrieval algorithms are improved. Complementary expansions of ground-based and aircraft-based CO₂ measurements in undersampled regions will similarly fill critical observational gaps in regions with large uncertainties and susceptibility to retrieval biases. Improvements to atmospheric CO₂ inversion systems, including reductions to systematic transport errors and improved error characterization, will be critical for refining top-down CO₂ budgets. And improved estimates of lateral carbon fluxes and managed land maps will refine estimates of agriculture, LULUCF, and waste emissions and removals.
Appendix A: TCCON sites

Table A1. TCCON sites used for evaluation of posterior CO$_2$ fields of the v10 OCO-2 MIP experiments.

<table>
<thead>
<tr>
<th>TCCON site</th>
<th>Country</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eureka</td>
<td>Canada</td>
<td>80.05° N</td>
<td>86.42° W</td>
<td>Strong et al. (2019)</td>
</tr>
<tr>
<td>Ny-Ålesund</td>
<td>Norway</td>
<td>78.9° N</td>
<td>11.9° E</td>
<td>Notholt et al. (2019b)</td>
</tr>
<tr>
<td>Sodankylä</td>
<td>Finland</td>
<td>67.4° N</td>
<td>26.6° E</td>
<td>Kivi et al. (2014)</td>
</tr>
<tr>
<td>East Trout Lake</td>
<td>Canada</td>
<td>54.4° N</td>
<td>105.0° W</td>
<td>Wunch et al. (2017a)</td>
</tr>
<tr>
<td>Bremen</td>
<td>Germany</td>
<td>53.10° N</td>
<td>8.85° E</td>
<td>Notholt et al. (2019a)</td>
</tr>
<tr>
<td>Karlsruhe</td>
<td>Germany</td>
<td>49.1° N</td>
<td>8.4° E</td>
<td>Hase et al. (2014)</td>
</tr>
<tr>
<td>Paris</td>
<td>France</td>
<td>48.8° N</td>
<td>2.4° E</td>
<td>Te et al. (2014)</td>
</tr>
<tr>
<td>Orléans</td>
<td>France</td>
<td>47.9° N</td>
<td>2.1° E</td>
<td>Warneke et al. (2019)</td>
</tr>
<tr>
<td>Garmisch</td>
<td>Germany</td>
<td>47.5° N</td>
<td>11.1° E</td>
<td>Sussmann and Rettinger (2018a)</td>
</tr>
<tr>
<td>Zugspitze</td>
<td>Germany</td>
<td>47.3° N</td>
<td>11.0° E</td>
<td>Sussmann and Rettinger (2018b)</td>
</tr>
<tr>
<td>Park Falls</td>
<td>USA</td>
<td>45.9° N</td>
<td>90.3° W</td>
<td>Wennberg et al. (2017)</td>
</tr>
<tr>
<td>Rikubetsu</td>
<td>Japan</td>
<td>43.5° N</td>
<td>143.8° E</td>
<td>Morino et al. (2014)</td>
</tr>
<tr>
<td>Lamont</td>
<td>USA</td>
<td>36.6° N</td>
<td>97.5° W</td>
<td>Wennberg et al. (2016b)</td>
</tr>
<tr>
<td>Anmeyondo</td>
<td>Korea</td>
<td>36.5° N</td>
<td>126.3° E</td>
<td>Goo et al. (2014)</td>
</tr>
<tr>
<td>Tsukuba</td>
<td>Japan</td>
<td>36.1° N</td>
<td>140.1° E</td>
<td>Morino et al. (2018a)</td>
</tr>
<tr>
<td>Nicosia</td>
<td>Cyprus</td>
<td>35.1° N</td>
<td>33.4° E</td>
<td>Petri et al. (2020)</td>
</tr>
<tr>
<td>Edwards</td>
<td>USA</td>
<td>34.2° N</td>
<td>118.2° W</td>
<td>Incat et al. (2016)</td>
</tr>
<tr>
<td>JPL</td>
<td>USA</td>
<td>34.2° N</td>
<td>118.2° W</td>
<td>Wennberg et al. (2016a)</td>
</tr>
<tr>
<td>Caltech</td>
<td>USA</td>
<td>34.1° N</td>
<td>118.1° W</td>
<td>Wennberg et al. (2014)</td>
</tr>
<tr>
<td>Saga</td>
<td>Japan</td>
<td>33.2° N</td>
<td>130.3° E</td>
<td>Kawakami et al. (2014)</td>
</tr>
<tr>
<td>Hefei</td>
<td>China</td>
<td>31.9° N</td>
<td>117.2° E</td>
<td>Liu et al. (2018)</td>
</tr>
<tr>
<td>Izaña</td>
<td>Spain</td>
<td>28.3° N</td>
<td>16.5° W</td>
<td>Blumenstock et al. (2017)</td>
</tr>
<tr>
<td>Burgos</td>
<td>Philippines</td>
<td>18.5° N</td>
<td>120.7° E</td>
<td>Morino et al. (2018b)</td>
</tr>
<tr>
<td>Manaos</td>
<td>Brazil</td>
<td>3.2° N</td>
<td>60.6° W</td>
<td>Dubey et al. (2014)</td>
</tr>
<tr>
<td>Ascension Island</td>
<td>UK</td>
<td>7.9° S</td>
<td>14.3° W</td>
<td>Feist et al. (2014)</td>
</tr>
<tr>
<td>Darwin</td>
<td>Australia</td>
<td>12.4° S</td>
<td>130.9° E</td>
<td>Griffith et al. (2014a)</td>
</tr>
<tr>
<td>Réunion Island</td>
<td>France</td>
<td>20.9° S</td>
<td>55.5° W</td>
<td>De Maziere et al. (2017)</td>
</tr>
<tr>
<td>Wollongong</td>
<td>Australia</td>
<td>34.4° S</td>
<td>150.9° E</td>
<td>Griffith et al. (2014b)</td>
</tr>
<tr>
<td>Lauder 125HR</td>
<td>New Zealand</td>
<td>45.0° S</td>
<td>169.7° E</td>
<td>Sherlock et al. (2014)</td>
</tr>
</tbody>
</table>

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/essd-15-963-2023-supplement.

Author contributions. The study was conceived of by DCr and designed by BB, DFB, SB, KWB, AC, FC, PC, NC, DCr, SF, LH, ARJ, JL, JBM, TO, PKP, BP, AS, CS and JRW. The v10 OCO-2 MIP experiments were performed by DFB, SB, MB, FC, NC, SC, DFB, ARJ, RJ, MSJ, DBAJ, JL, ZL, SM, SMM, MR, AS, AZM and NZ. Lateral flux estimates were performed by FC, PC, ZD, HT and YY. FF emission estimates were performed by TO and SB. TCCON data were collected by NMD, MKD, OEG, BH, RK, IM, JN, YSO, HO, CP, KS, KS, YT, VAV, MV, TW and DW. Evaluation of v10 OCO-2 MIP against co-samples was performed by BB, SB, AC and ARJ. BB wrote the paper and prepared the figures, with contributions from all co-authors.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Earth System Science Data. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (grant no. 80NM0018D004). The v10 OCO-2 MIP activity was supported by the NASA OCO science team program. Brendan Byrne and Junjie Liu were supported by the NASA OCO-2
and OCO-3 science team program (grant no. NNH17ZDA001N-OCO2). Andrew R. Jacobson, Andrew Schuh and David F. Baker were funded by NASA (grant no. 80NSSC21K1080). Andrew Schuh was also supported by NASA (grant no. NNX15AG93G). Sean Crowell acknowledges support from the NASA OCO science team (grant no. 80NSSC21K1077). The research of Noel Cressie, Andrew Zammit-Mangion, and Michael Bertolacci was supported by the Australian Research Council Discovery Project (grant no. DP190100180) and by NASA ROSES (grant no. 20-OCOST20-0004). Andrew Zammit-Mangion is also supported by the Australian Research Council Discovery Early Career Research Award (grant no. DE180100203). The contribution of Abhishek Chatterjee was supported by NASA ROSES (grant nos. 80NSSC20K0006 and 80NSSC21K1068). Matthew S. Johnson acknowledges the internal funding from NASA’s Earth Science Research and Analysis Program. The contributions of Frédéric Chevallier and Marine Remaud were supported by the Copernicus Atmosphere Monitoring Service and implemented by the European Centre for Medium-Range Weather Forecasts on behalf of the European Commission (grant no. CAMST3) and by the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement number 958927 (prototype system for a Copernicus CO₂ service). It was granted access to the HPC resources of TGCC under the allocation A01I0102201 made by GENCI. Sha Feng at PNNL is supported by NASA Carbon Monitoring Program (grant no. 80HQT21T0069). The Pacific Northwest National Laboratory is operated by Battelle Memorial Institute under contract DE-AC05-76RL01830. Ning Zeng acknowledges support from NOAA (grant no. NA180AR4310266) and NASA (grant no. 80NSSC21K10908). Prabir K. Patra is partly supported by the Environment Research and Technology Development Fund (grant no. JPMEEF21S20800) of the Environmental Restoration and Conservation Agency of Japan. Scot M. Miller was supported by the NASA grants 80NSSC18K0976 and 80NSSC21K1073. Nicholas M. Deutscher was funded by the ARC Future Fellowship (grant no. FT180100327).

The TCCON Nicosia site has received additional support from the European Union’s Horizon 2020 Research and Innovation Programme (grant no. 856612/EMME-CARE), the Cyprus Government and the University of Bremen. The TCCON Ammyeondo site has been funded by the Korea Meteorological Administration Research and Development Program, Developing Technology for Integrated Climate Change Monitoring and Analysis, under grant KMA2018-00320. TCCON sites at Tsukuba, Rikubetsu and Burgos are supported in part by the GOSAT series project. Burgos is supported in part by the Energy Development Corp., Philippines. The Eureka TCCON measurements were made at the Polar Environment Atmospheric Research Laboratory (PEARL) by the Canadian Network for the Detection of Atmospheric Change (CANDAC), primarily supported by the Natural Sciences and Engineering Research Council of Canada, Environment and Climate Change Canada, and the Canadian Space Agency. The TCCON site at Réunion Island has been operated by the Royal Belgian Institute for Space Aeronomy with financial support since 2014 by the EU project ICOS-Inwire and the ministerial decree for ICOS (FR/35/IC1 to FR/35/C6) and local activities supported by LACY/UMR8105 and by OSU-R/UMS3365 – Université de La Réunion. Darwin and Wollongong TCCON stations are supported by ARC grants DP160100598, LE0668470, DP140101552, DP110103118 and DP0879468 and NASA grants NAG512247 and NNG05GD07G.

We thank Robert J. Andres for providing uncertainty estimates for CDIAC fossil fuel emission estimates. We thank the data providers of the SeaFlux ensemble for making their pCO₂-based air–sea CO₂ fluxes publicly available. We are grateful for the leadership of Annemarie Eldering and Mike Gunson of the OCO-2 mission, whose hard work has made this dataset possible.

Financial support. This research has been supported by the European Commission, Horizon 2020 Framework Programme (CoCO2 grant no. 958927 856612/EMME-CARE)) and Copernicus Atmosphere Monitoring Service (grant no. CAMST3), the Australian Research Council (grant nos. DP190100180, DE180100203, DP160100598, LE0668470, DP140101552, DP110103118, DP0879468 and FT180100327), the Environmental Restoration and Conservation Agency (grant no. JPMEEF21S20800), the Korea Meteorological Administration (grant no. KMA2018-00320), the National Aeronautics and Space Administration (grant nos. 20-OCOST20-0004, 80NSSC18K0908, 80NSSC18K0976, 80NSSC20K0006, 80NSSC21K1068, 80NSSC21K1073, 80NSSC21K1077, 80NSSC21K1080, 80HQT21T0069, NAG512247, NNG05GD07G, NNH17ZDA001N-OCO2 and NNX15AG93G), and the National Oceanic and Atmospheric Administration (grant no. NA180AR4310266).

Review statement. This paper was edited by David Carlson and reviewed by two anonymous referees.

References


Caspersen, J. P., Pacala, S. W., Jenkins, J. C., Hurt, G. C., Moorcroft, P. R., and Birdsey, R. A.: Contributions of land-use history


Harris, N. L., Brown, S., Hagen, S. C., Saatchi, S. S., Petrova, S., Salas, W., Hansen, M. C., Potapov, P. V., and Lotosch, A.: Baseline map of carbon emissions from


Oda, T. and Maksyutov, S.: ODIAC Fossil Fuel CO2 Emissions Dataset (Version name 1: ODIACYYYY or ODIACYYYYa), Center for Global Environmental Research,


Batchelor, R., and Lindenmaier, R.: TCCON data

vertical profiles of atmospheric CO
S.: Weak northern and strong tropical land carbon uptake from
N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Lange-


Tohjima, Y., Mukai, H., Machida, T., Nojiri, Y., and Gloor, M.: First measurements of the latitudinal atmospheric O2 and CO2

Plass-Duelmer, C., Ramonet, M., Ramos, R., Reyes-Sanchez, E., Richardson, S., Riris, H., Rivas, P. P., Ryerson, T., Saito, K., Sargent, M., Sasaki, M., Sawa, Y., Say, D., Scheeren, B., Schuck, T., Schumacher, M., Seifert, T., Sha, M. K., Shep-


Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Pe-


