Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2019GB006393

Key Points:
- Top-down and bottom-up estimates of net land-atmosphere CO_2 fluxes agree well globally but show important mismatches at regional scales
- Regional mismatches are dominated by differences between inversions and interannual variability in CO_2 fluxes
- Mismatches between top-down and bottom-up data sets are explained by sensitivity to climate and by uncertainty in land use change forcing

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Bastos,
ana.bastos@lmu.de

Citation:
Bastos, A., O'Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., et al. (2020), Sources of uncertainty in regional and global terrestrial CO_2 exchange estimates. *Global Biogeochemical Cycles*, *34*, e2019GB006393. https://doi.org/10.1029/2019GB006393

Received 8 AUG 2019
Accepted 24 JAN 2020
Accepted article online 7 FEB 2020

©2020. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Sources of Uncertainty in Regional and Global Terrestrial CO_2 Exchange Estimates

A. Bastos[1], M. O'Sullivan[2], P. Ciais[3], D. Makowski[4,5], S. Sitch[6], P. Friedlingstein[2], F. Chevallier[3], C. Rödenbeck[7], J. Pongratz[1,8], I. T. Luijkx[9], P. K. Patra[10], P. Peylin[3], J. G. Canadell[11], R. Lauerwald[12], W. Li[3,13], N. E. Smith[9], W. Peters[9,14], D. S. Goll[15], A.K. Jain[16], E. Kato[17], S. Lienert[18], D. L. Lombardozzi[19], V. Haverd[20], J. E. M. S. Nabel[8], B. Poulter[21], H. Tian[22], A. P. Walker[23], and S. Zaehle[7]

[1]Department of Geography, Ludwig-Maximilians Universität, München, Germany, [2]College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK, [3]Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, UMR8212, Gif-sur-Yvette, France, [4]University Paris-Saclay, AgroParisTech, INRAE, UMR 211, Thiverval-Grignon, France, [5]CIRED, Nogent-sur-Marne, France, [6]College of Life and Environmental Sciences, University of Exeter, Exeter, UK, [7]Max Planck Institute for Biogeochemistry, Jena, Germany, [8]Max Planck Institute for Meteorology, Hamburg, Germany, [9]Department of Meteorology and Air Quality, Wageningen University and Research, Wageningen, Netherlands, [10]Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan, [11]CSIRO Oceans and Atmosphere, Canberra, ACT, Australia, [12]Department of Geoscience, Environment and Society, Université Libre de Bruxelles, Bruxelles, Belgium, [13]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China, [14]Centre for Isotope Research, University of Groningen, Groningen, Netherlands, [15]Department of Geography, University of Augsburg, Augsburg, Germany, [16]Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, IL, USA, [17]Institute of Applied Energy, Minato, Japan, [18]Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland, [19]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [20]CNRM, Université de Toulouse, Mto-France, CNRS, Toulouse, France, [21]Biospheric Sciences Lab, NASA, Greenbelt, MD, USA, [22]International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA, [23]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA

Abstract The Global Carbon Budget 2018 (GCB2018) estimated by the atmospheric CO_2 growth rate, fossil fuel emissions, and modeled (bottom-up) land and ocean fluxes cannot be fully closed, leading to a "budget imbalance," highlighting uncertainties in GCB components. However, no systematic analysis has been performed on which regions or processes contribute to this term. To obtain deeper insight on the sources of uncertainty in global and regional carbon budgets, we analyzed differences in Net Biome Productivity (NBP) for all possible combinations of bottom-up and top-down data sets in GCB2018: (i) 16 dynamic global vegetation models (DGVMs), and (ii) 5 atmospheric inversions that match the atmospheric CO_2 growth rate. We find that the global mismatch between the two ensembles matches well the GCB2018 budget imbalance, with Brazil, Southeast Asia, and Oceania as the largest contributors. Differences between DGVMs dominate global mismatches, while at regional scale differences between inversions contribute the most to uncertainty. At both global and regional scales, disagreement on NBP interannual variability between the two approaches explains a large fraction of differences. We attribute this mismatch to distinct responses to El Niño–Southern Oscillation variability between DGVMs and inversions and to uncertainties in land use change emissions, especially in South America and Southeast Asia. We identify key needs to reduce uncertainty in carbon budgets: reducing uncertainty in atmospheric inversions (e.g., through more observations in the tropics) and in land use change fluxes, including more land use processes and evaluating land use transitions (e.g., using high-resolution remote-sensing), and, finally, improving tropical hydroecological processes and fire representation within DGVMs.

1. Introduction

The United Nations Framework Convention on Climate Change Paris Agreement from 2015 (UNFCCC, 2015) has the goal to limit the increase in global average temperature well below 2 °C above preindustrial

levels, and pursue efforts not to exceed the target of 1.5 °C above preindustrial levels. To achieve this overarching goal, the agreement calls for "a balance between anthropogenic emissions by sources and removals by sinks of greenhouse gases in the second half of this century."

On the policy side, the problem is expected to be tackled at the national level through Nationally Determined Contributions, that is, mitigation policies defined by countries (UNFCCC, 2015). The collective progress toward the overall goal of the agreement is to be assessed regularly in the "global stocktake" process to support the update of Nationally Determined Contributions according to the best available science. However, the ultimate goal of stabilizing global mean surface temperature requires evaluating common progress at the global scale, as well as the evolution of natural sinks. It is therefore crucial that bottom-up estimates of CO_2 fluxes (e.g., inventories and models) provide accurate estimates that are consistent with the global atmospheric greenhouse gas mole fractions.

Global carbon (C) budgets for the anthropogenic perturbation have been estimated in the successive Intergovernmental Panel on Climate Change (IPCC) assessment reports (IPCC, 2001, 2007, 2013) and have been revised and updated in the Global Carbon Budget (GCB) by the Global Carbon Project almost every year since 2005 (Le Queré et al., 2009; Le Quéré, Andrew, Friedlingstein, Sitch, Pongratz, et al., 2018; Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b).

Fossil fuel and cement production emissions (E_{FF}) can be estimated from historical energy statistics (Boden et al., 2017; UNFCCC, 2018) with a 1 standard deviation (1σ) uncertainty of 5–11% (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b; Quilcaille et al., 2018). The fluxes from changes in land use and management (F_{LUC}) cannot be directly measured and are estimated in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) by bookkeeping models (Hansis et al., 2015; Houghton & Nassikas, 2017), with a reported uncertainty of 0.7 Pg C yr^{-1} (1σ) for decadal average F_{LUC} over the industrial era. Global carbon uptake by the ocean and land can also not be directly measured, and are estimated in the GCB by process-based models for the land, and process- and data-driven models for the ocean (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b; Sitch et al., 2013).

In the Global Carbon Budget 2018 (GCB2018; Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b), the net annual balance between the anthropogenic sources and the sinks of CO_2 from process-based models does not exactly match the accurately observed atmospheric CO_2 growth rate (G_{ATM}; Dlugokencky & Tans, 2018). The residual flux resulting from this gap, the "budget imbalance," can be interpreted as a measure of the limitations of the data sets used and of the imperfect process understanding by the modeling community (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b). Since the budget imbalance does not show a significant trend since the 1960s but high year-to-year variability, Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) proposed that errors in the land and ocean sinks explain most of the imbalance term, but uncertainty in F_{LUC} is also known to be high (Arneth et al., 2017; Piao et al., 2018). In GCB2018, estimates of ocean and land net CO_2 fluxes from process-based models were compared with those from four observation-based atmospheric inversions (Chevallier et al., 2005; Rödenbeck et al., 2003; Saeki & Patra, 2017; van der Laan-Luijkx et al., 2017) for the globe and over three latitudinal bands, where large differences between process-based models and inversions, but also between inversions were found. A recent study (Gaubert et al., 2019) further compared 10 different inversion systems, including the four mentioned above, showing an overall fair match with aircraft CO_2 measurements and varying consistency (<10% on 3-year mean varying with the inversion) with G_{ATM}. However, their results showed that differences in fossil fuel emission priors used affected the ocean-land partitioning, setting limits to the accuracy of inversions for quantifying regional surface-atmosphere CO_2 fluxes. Even though inversions were adjusted for differences in E_{FF} priors in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b), large disparities were still found for regional fluxes estimated by the four atmospheric inversions, especially the balance between the Northern Hemisphere and the tropics.

Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) proposed that regional-level analyses may uncover sources of errors in the inferred fluxes. For example, comparisons of net CO_2 surface flux estimates from atmospheric inversions and dynamic global vegetation models (DGVMs) allowed assessing in detail the response of the tropical net CO_2 exchange to the 2015–2016 El Niño event, and evaluating the DGVMs' ability to simulate the anomalies in carbon fluxes in 2015/16 (Bastos et al., 2018; Gloor et al., 2018; Rödenbeck et al., 2018; van Schaik et al., 2018). Likewise, the studies produced in the framework of the REgional Carbon Cycle Assessment and Processes (RECCAP; Canadell et al., 2012) provided a first evaluation of global and regional budgets, budget component fluxes and uncertainties. RECCAP produced multiple global syntheses

of regional CO_2 fluxes between 1990 and 2009 and their driving processes (Andres et al., 2012; Ciais et al., 2014; Chen et al., 2013; Houghton et al., 2012; Khatiwala et al., 2013; Peters et al., 2012; Peylin et al., 2013; Sitch et al., 2013; Wanninkhof et al., 2013) as well as independent budgets for nine land, four ocean and coastal regions.

In this study, we take a closer look at the regional mismatch between the top-down and bottom-up data sets in GCB2018 (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b) to identify the sources of uncertainty in the terrestrial components, specifically to:

(i) identify the regions driving uncertainties in global budget estimates;
(ii) quantify the relative contribution of uncertainty in net CO_2 surface flux estimates from inversions (top-down) and from DGVMs (bottom-up);
(iii) identify those processes contributing the most to uncertainty in global and regional budgets.

We focus on a group of 18 land regions from Tian et al. (2018), which are consistent with the previous nine RECCAP large regions and with ecoclimatic specificities, but follow political borders. Evaluating the agreement between top-down and bottom-up estimates of CO_2 fluxes over such regions allows assessing the confidence and limitations of the state-of-the-art science to provide information on regional budgets, relevant for the global stocktaking process yet with traceable consistency with the G_{ATM}.

In this study, we rely on net CO_2 exchange from inversions, considered as uncertain estimates of surface CO_2 fluxes consistent with G_{ATM} and compare inversions' estimates to those from the 16 DGVMs in GCB2018. Since Gaubert et al. (2019) recently analyzed the sources of uncertainty between inversions, here we focus on predictors that can explain the differences between DGVMs and inversions, or differences between individual DGVMs. By statistically modeling the mismatch between inversions and DGVMs, we are able to attribute uncertainty in the global budgets to a few key regions mainly in the tropics, and to identify those processes contributing the most to regional and global uncertainties.

2. Data

2.1. Atmospheric Inversions

Atmospheric inversions use an optimization process by which atmospheric CO_2 mole fraction measurements are used to constrain a priori estimates of the spatial distribution of surface CO_2 fluxes using an atmospheric transport model. In the optimization, information about errors of measurements and of priors, errors in the transport model, and the spatiotemporal structure of fluxes is also included (Peylin et al., 2013).

The four inversion systems used in this study (Table 1) are all based on in situ CO_2 observation measurements but differ in several aspects, such as the observational data assimilated and period covered, the transport model and prior fluxes used, the model grid, the spatiotemporal a priori correlation structure, as well as in the fossil fuel data sets used. A complete description of the different atmospheric inversions can be found in Table A3 in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b). Here we use two versions of CarboScope (Rödenbeck et al., 2003), both covering a period longer than 30 years but with variable number of assimilated sites: s76 (1976–2017, eight stations) and s85 (1985–2017, 21 stations). Thus, we have used in total five inversion data sets (CAMS, two CarboScope versions, MIROC, and CarbonTracker Europe; Table 1).

Because they are based on CO_2 concentration measurements, atmospheric inversions estimate the net surface-atmosphere CO_2 fluxes including both the natural (fires, storms, pests, and diseases) and anthropogenic disturbance terms, the subsequent recovery, as well as the carbon taken up from the atmosphere over land but then passed on to the oceans through freshwaters, estuaries, and coastal areas (Hartmann et al., 2009; Mayorga et al., 2010; Regnier et al., 2013).

As in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) and Peylin et al. (2013), we adjusted the ocean and land fluxes for differences in fossil fuel emission (E_{FF}) priors over large latitudinal bands. As reference E_{FF}, we chose the data used by CAMS, the Emission Database for Global Atmospheric Research (Olivier et al., 2017) scaled to the Carbon Dioxide Information Analysis Center (Marland et al., 2008) estimates. However, this is only a first-order correction as the biases in E_{FF} not only affect the flux estimation of the region in question but also the neighboring regions (Saeki & Patra, 2017). The inversion surface fluxes were then remapped to a regular $1° \times 1°$ latitude/longitude grid and then aggregated to the 18 land regions.

Table 1
Top-Down and Bottom-Up Data Sets Used in This Study, Including References and Period Covered

Data set/model	Reference	Period	Fire	N-dep	SC	WH
	Inversions (top-down)					
Data set						
Copernicus Atmosphere Monitoring Service (CAMS)	Chevallier et al. (2005)	1979–2017				
CarboScope s76	Rödenbeck et al. (2003)	1976–2017				
CarboScope s85		1985–2017				
MIROC	Patra et al. (2018)	1997–2017				
CarbonTracker Europe (CTE)	van der Laan-Luijkx et al. (2017)	2001–2017				
	DGVMs (bottom-up)					
Model						
CABLE-POP	Haverd et al. (2018)	1979–2017	N	Y	Y	Y
CLASS-CTEM	Melton and Arora (2016)		Y	Y	N	N
CLM5.0	Oleson et al. (2013)		Y	N	Y	Y
DLEM	Tian et al. (2015)		N	Y	N	Y
ISAM	Meiyappan et al. (2015)		N	Y	N	Y
JSBACH	Mauritsen et al. (2018)		Y	Y	Y	Y
JULES	Clark et al. (2011)		N	N	N	N
LPJ	Poulter et al. (2011)		Y	N	Y	Y
LPJ-GUESS	Smith et al. (2014)		Y	Y	Y	Y
LPX-Bern	Lienert and Joos (2018)		Y	Y	N	N
OCN	Zaehle et al. (2010)		N	Y	N	Y
ORCHIDEE	Krinner et al. (2005)		N	N	N	Y
ORCHIDEE-CNP	Goll et al. (2017)		N	Y	N	N
SDGVM	Walker et al. (2017)		Y	Y	N	N
SURFEX	Joetzjer et al. (2015)		Y	N	N	N
VISIT	Kato et al. (2013)		Y	N	Y	Y

Note. For the DGVMs, we indicate whether they simulate fires, nitrogen deposition (N-dep), shifting cultivation (SC), and wood harvest (WH), and in some models further include irrigation and nitrogen fertilization. A complete description of the processes in each model can be found in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b).

Inversions determine total CO_2 fluxes between the surface and the atmosphere whereas the GCB approach with land and ocean biogeochemical models determines the anthropogenic budget of CO_2. The difference between total and anthropogenic CO_2 fluxes is that there is a background preindustrial uptake of CO_2 on land that sustains carbon export from soils to rivers and to the ocean, where compensatory outgassing of CO_2 occurs. This requires an adjustment of inversion CO_2 fluxes to transform them into anthropogenic CO_2 fluxes, as performed in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) over latitudinal bands.

Based on the data-driven estimates of fluvial exports of organic (Mayorga et al., 2010) and inorganic carbon (Hartmann et al., 2009) to the coast, Zscheischler et al. (2017) produced a spatially explicit data set of climatological land-ocean carbon transfers at $1° \times 1°$ latitude/longitude resolution and includes the fluxes from dissolved inorganic carbon from atmospheric origin and from weathering and dissolved and particulate organic carbon (DOC and POC). In this study, the DOC and POC exports of this data set were rescaled per basin to match the estimates of Resplandy et al. (2018). After aggregating these rescaled estimates to the 18 land regions, we subtracted the fluvial carbon exports from the regionally aggregated inversion net surface CO_2 fluxes, to calculate regional net biospheric production (NBP) that can be compared with the DGVM estimates. These data are available in (Bastos, 2019).

2.2. DGVMs

DGVMs simulate water, energy and biogeochemical exchanges between the surface and the atmosphere through ecosystems activity, including growth, turnover and decomposition of vegetation, and soil carbon

processes. DGVMs simulate the response of ecosystems to changes in environmental conditions (increasing CO_2, changing climate, and altered nitrogen deposition) and land use activity.

Here we use a set of 16 DGVMs from the TRENDY-v7 intercomparison (Sitch et al., 2015) that contributed to GCB2018. In the simulations for GCB2018 (Simulation S3), all 16 models are forced with: (i) observed climate from the Climate Research Unit (Harris et al., 2014) and the Japanese 55-year Reanalysis (Kobayashi et al., 2015) data sets, the CRU-JRA reanalysis, following the methodology in Viovy (2016); (ii) global CO_2 concentration from NOAA/ESRL (Dlugokencky & Tans, 2018); (iii) land use change transitions and land management fields from the Land Use Harmonization (LUH2v2.1h) from Hurtt et al., 2011 (2011, 2017), based on HYDE 3.1 (Klein Goldewijk et al., 2011); and (iv) gridded data of nitrogen (N) deposition when N cycling is simulated by models.

Some models simulate fire emissions (E_{Fire}) and nutrient cycling (nitrogen in ten models, and nitrogen and phosphorus in one model), including the effects of N deposition (Table 1). While all DGVMs simulate the fluxes resulting from forest clearing, pasture and crop conversion, abandonment and regrowth and crop harvest, they implement them using different assumptions about the areas being converted (e.g., gross versus net conversion). The way DGVMs simulate management practices and the fate of released carbon also varies between models. Ten models simulate wood harvest and six include shifting cultivation, roughly corresponding to gross transitions (Table 1). Only few DGVMs simulate crop fertilization or irrigation. To calculate F_{LUC}, the S3 simulation is compared to the S2 simulation, which is forced with CO_2 and climate changes only, and keeping a fixed land cover map in 1700 (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b).

For F_{LUC} estimated in this way, fluxes resulting from environmental change over managed lands as compared to intact vegetation are included (Pongratz et al., 2014). Due to the effect of CO_2 fertilization, intact vegetation would at present provide a slightly stronger sink, which implies that emissions from, for example, deforestation are now slightly higher than if compared to the preindustrial state. This term is referred to as loss of additional sink capacity (LASC) as proposed by Pongratz et al. (2014). In TRENDY-v7, an additional simulation was performed where preindustrial CO_2 and climate were kept fixed, and only land use changes and management were allowed to change between 1700 and 2017 (S4). F_{LUC} estimated in this way does not include the LASC term and is therefore more comparable in terms of processes to the ones estimated by the bookkeeping methods, except for the values of C densities used.

Outputs of monthly NBP from all three simulations (S2, S3, and S4) were first resampled to a common $1° \times 1°$ latitude/longitude grid, and then aggregated for each region. The regional fluxes from S2 and S3 simulations are available from (Bastos, 2019).

2.3. Bookkeeping Models

Bookkeeping models track changes in above- and below-ground biomass carbon densities resulting from land use change processes such as deforestation/afforestation, cropland or pasture expansion, wood harvest, shifting cultivation, and forest regrowth after land abandonment. In GCB2018, two bookkeeping models that estimate F_{LUC} at the global scale have been considered: the bookkeeping model of Houghton and Nassikas (2017), referred as HN2017 henceforth, and the "Bookkeeping of Land Use Emissions" model (BLUE) described in Hansis et al. (2015). To estimate F_{LUC}, bookkeeping models first calculate the changes in biomass, soil and the atmosphere carbon pools resulting from a given transition following specific response curves. The resulting fluxes for a given year can be calculated as the difference in carbon stocks between two consecutive years. The carbon density values for above- and below-ground biomass used by both models are based on recent history (ca. 1980s) measurements. This means that the C densities used have an implicit transient effect of increasing CO_2 and climate, leading to somewhat higher CO_2 emissions than S4 by DGVMs.

Even though based on similar principles, the two models differ in their mathematical formulation, spatial implementation, assumptions made about LUC transitions, processes included, forcing data used and, consequently, in the sources of uncertainty.

The HN2017 model covers the period of 1700–2015 and is based on regional statistics from the Food and Agricultural Organization (FAO) on changes in the areas of croplands and pastures since 1961 and changes in the areas of forests and other land since 1990 (FAO, 2015; FAOSTAT, 2015), while BLUE used the same data set as the DGVMs, that is, LUH2v2.1h. In the HN2017 model, calculations are performed at the country

scale, and in the recent version, the model now incorporates country specific C densities. In this version, the model does not include shifting cultivation. From 1997 onward, peat fires (Giglio et al., 2016) and emissions from peat drainage (Hooijer et al., 2010) are added in Southeast Asia. The fluxes resulting from fire suppression are considered, but only in the United States. The HN2017 model allocates pasture preferentially to grasslands. This may result in lower CO_2 emissions by reducing deforestation rates (Reick et al., 2010).

The BLUE model uses a spatially explicit modeling scheme to calculate F_{LUC} on a pixel basis between 850 and 2017. It relies on biome-specific C densities and exponential response curves to track the changes in soil and biomass carbon pools following land use conversion or due to management. New cropland and pasture are taken proportionally from natural vegetation types and rangelands clear the natural vegetation for forest areas and degrade other natural land (Hansis et al., 2015; Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b). BLUE further includes shifting cultivation. In the GCB2018 version, BLUE used the same data set as several of the DGVMs, that is, the LUH2v2.1h, to calculate F_{LUC} at 0.25° × 0.25° latitude/longitude resolution. Because BLUE uses a spatially explicit framework and the same LUC forcing as DGVMs, we use it as a primary basis for comparison with process-based model estimates of F_{LUC}. We then compare BLUE with HN2017 to evaluate the potential contribution of the forcing used to estimate F_{LUC} by DGVMs and BLUE to the mismatch between DGVMs and inversions.

2.4. Additional Data

In order to evaluate possible sources of uncertainty, we considered additional data sets to explain the differences between DGVM and inversion fluxes (see section 2.5): the Oceanic Niño Index (ONI) from NOAA's Climate Prediction Center (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v4.shtml) from December to March; annual average total water storage (TWS) over each of the 18 regions from the Gravity Recovery and Climate Experiment (GRACE) reconstruction (Humphrey et al., 2018); satellite-based land cover maps from the European Space Agency's Climate Change Initiative Land-Cover (LC-CCI) (ESA, 2017) from 1992–2015 aggregated from 250 m × 250 m to 1° × 1° latitude/longitude resolution; carbon emissions data from the Global Fire Emissions Database version 4.1s (GFED4.1s, https://doi.org/10.3334/ORNLDAAC/1293) available from 1997–2017, described in Randerson et al. (2017); and burned area from ESA-CCI Fire-CCI v5.1 (Fire-CCI; Chuvieco et al., 2018), covering the period 2001–2017.

2.5. Methods

The goal here is to quantify the mismatch between NBP estimated by inversions and DGVMs. As discussed in section 2.1, inversions differ in several aspects between each other. Likewise, DGVMs show variable skill in simulating fundamental processes of the terrestrial C cycle (see Figure B2 in Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b) and do not all simulate the same processes (Table 1). Some of the differences across data sets may arise from complex interactions between processes within models, and not be traceable to a particular process.

We define D_{ikj} as the difference in year j ($j = 1, \ldots, J$) between NBP estimates of the kth inversion ($k = 1, \ldots, n_j$, with $n_j \leq 5$, depending on the year j) available over that year and the ith DGVM ($i = 1, \ldots, 16$).

$$D_{ikj} = DGVM_{ij} - INV_{kj} \tag{1}$$

For a given year, the NBP values estimated by the inversions (INV_{kj}) can be then considered as a sample from a distribution of surface fluxes compatible with G_{ATM}, with the size of this sample changing over the study period (1979–2017) as more inversions become available. For each year we can generate a set of $n_j \times 16$ values of D for the globe (D_{Globe}), as well as for each of the 18 regions (D_{reg}).

For each (inversion, DGVM) pair, the D_{Globe} can be decomposed into the 18 region components:

$$D_{Globe_{ikj}} = \sum_{r=1}^{18} D_{Rreg_{rikj}} \tag{2}$$

We can, therefore, quantify the contribution of each region to the variance (V) of D_{Globe} as

$$V(D_{Globe_{ik}}) = \sum_{r=1}^{18} V(D_{Rreg_{rik}}) + 2\sum_{r=1}^{18}\sum_{p=1}^{18} \mathrm{Cov}\ (D_{Rreg_{rik}} D_{Rreg_{pik}}) \tag{3}$$

The contribution of each region r to the variance of D_{Globe} is therefore the quotient between the sum of the regional variance and covariance terms in r and the variance of D_{Globe}, for each (inversion, DGVM) pair.

To identify possible drivers of D, we test several statistical models using likely predictors for the mismatch between inversions and DGVMs. Linear mixed effects (LME) models allow modeling D as a function of fixed effects (e.g., El Niño/La Niña cycles), but also of random effects describing several sources of variability (DGVMs, years, inversions). We start by fitting to D_{ikj} the following simple LME model without any predictor region by region (the lme function from the R package lme4; Bates et al., 2014) described in equation (4) (LME_{RE}):

$$D_{ikj} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ikj} \tag{4}$$

where μ is the average mismatch (i.e., D averaged over all models, years, and inversions) over the considered region. The random effects α_i, β_j and γ_{ij} (all with zero mean and variance V_α, V_β, and V_γ, respectively) describe the variability of D across the DGVMs with index i, the between-year variability of D with index j, and the interaction between models and years on D, respectively. The interaction term accounts for the possibility that some models may have larger D in some years, and smaller D in other years. The ϵ_{ikj} term describes the variability across the five inversions and their interactions with year and DGVMs with index k and variance V_ϵ. For each region, a set of LME models with the possible combinations of one, two or all three random effects are fit to D, the model that best explains D for each region is retained, by choosing the one with lowest Akaike Information Criterion and with significant fit (p value < 0.05). We then expand the LME in equation (4) in order to explain part of the variance of D using different sets of predictors. We test four hypotheses (equation (5)) that are supported by literature and for which data are available. We define a set of LME models (equation (5)) (LME_{FE}) with one or more predictors (X_a, X_b, X_c, X_d) and associated coefficients (c_a, c_b, c_c, c_d) corresponding to the following hypotheses:

a. D can be explained by a linear time trend, for example, if models under- or overestimate the sensitivity to a process with a strong trend component such as CO_2 increase (Graven et al., 2013; Thomas et al., 2016);
b. Since D corresponds to the difference between top-down and bottom-up estimates of NBP, and NBP includes F_{LUC}, which is highly uncertain (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b; Piao et al., 2018), D might be explained by errors in F_{LUC};
c. Differences can be due to too strong response of DGVMs to ENSO (Bastos et al., 2018);
d. In fire-dominated regions, D may be explained by the fact that some DGVMs do not simulate fires (Table 1), or because those that do still show limited ability to correctly represent spatial and interannual variability in burned area and fire emissions (Li et al., 2014; Poulter et al., 2015; Yue et al., 2014, 2015).

The full model includes the four predictors and is expressed as

$$D_{ikj} = c_a \times X_{aj} + c_b \times X_{bj} + c_c \times X_{cj} + c_d \times X_{dj} + \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ikj} \tag{5}$$

Because the predictors have different units, and variability across regions may also vary significantly for F_{LUC} and E_{fire}, we centered and standardized these variables (i.e., the mean was subtracted and the result divided by the standard deviation). As done for the random effects, we define multiple models with one, two, three, or four predictors as fixed effects, and choose the best model fit (lowest Akaike information criterion and significant fit). We fit the LME_{FE} separately for 1997–2017 when including fire emissions as a predictor.

When any of these variables is found to provide the best fit, we explore how the predictors may contribute to explain the differences. In the case of F_{LUC} we compare the estimates from DGVMs with those of BLUE and HN2017 and evaluate the role of the forcing LUC to explain D in that region. As ENSO effects on NBP are mainly related to changes in temperature and water availability in the tropics (warm/dry during El Niño and cool/wet during La Niña) we evaluate the sensitivity of fluxes to soil water availability and temperature by performing a multiple linear regression of spatially aggregated fluxes with regionally averaged annual temperature from CRU-JRA and water availability. Water availability was estimated by annual TWS from GRACE for inversions and simulated soil moisture for DGVMs. Where E_{fire} is found to contribute to D, we compare simulated emissions from fire as well as burned area with the reference data sets.

We further group DGVMs by "flavors," corresponding to whether they represent (or not) the processes highlighted in Table 1: fire, N deposition, shifting cultivation and wood harvest. We evaluated whether

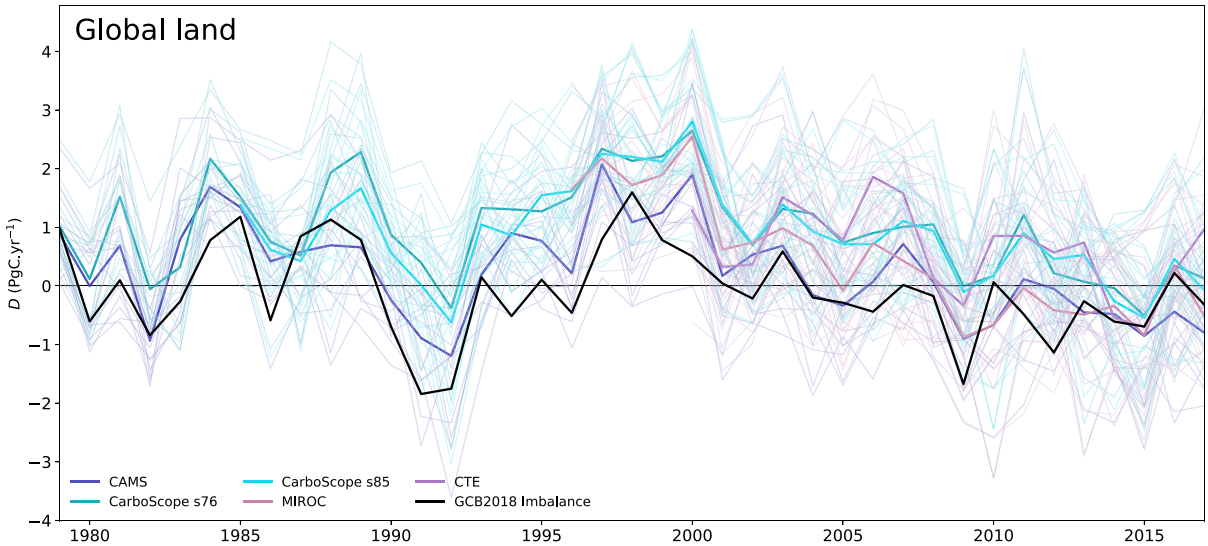

Figure 1. Mismatch between DGVMs and inversions for the global land sink estimates, calculated following equation (1). The thin colored lines correspond to *D* for each inversion and each model, with the colors indicating the respective inversion (CAMS in dark blue; the two CarboScope inversions s76 and s85 in dark and light cyan, respectively; MIROC in pink; and CarbonTracker Europe [CTE] in purple). The bold lines show the multimodel ensemble mean of *D* for each inversion. The budget imbalance calculated by Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al. (2018b) is shown in black (inverted sign to match the sign of equation (1)) for comparison.

the different groups of DGVMs have significantly lower biases (μ) than the set of 16 DGVMs at global and regional scale (supporting information Table S1), and additionally add the process representation as a predictor in the LME_{FE} fit (Figure S7).

3. Results

3.1. Global Differences and the "Budget Imbalance"

The values of G_{ATM} reconstructed from the inversions fluxes on average match those from GCB2018 (Figure S1), with differences of −0.04±0.41, 0.11±0.36, 0.14±0.36, 0.29±0.42, and 0.00±0.48 Pg C yr^{-1}, for CAMS, CarboScope s76, CarboScope s85, MIROC, and CarbonTracker Europe, respectively, over the period covered by each inversion. For comparison, the budget imbalance term in GCB2018 for 2008–2017 was 0.5 Pg C yr^{-1}.

The ensemble of D_{Globe} calculated following equation (1) for all (inversion, DGVM) pairs is shown in Figure 1. Over the 1979–2017 period, D_{Globe} averaged across all DGVMs for each inversion (bold lines) show strong multiannual fluctuations. From 1983–1989 (except 1982) and over the late 1990s, D_{Globe} was strongly positive, which indicates stronger NBP from DGVMs compared to inversions; following the year 2000, D_{Globe} showed a decreasing trend. These multiannual tendencies are punctuated by years of strong peaks, for example, 1982, 1991/1992, 2009/2010, and 2015 (negative D_{Globe}) and 1984, 1988/1989, 2000, and 2011 (positive). The trough in 1991/1992 suggests that DGVMs miss the large abnormal sink deduced from G_{ATM}, which is a known feature from all previous budgets (Le Queré et al., 2009; Le Quéré, Andrew, Friedlingstein, Sitch, Pongratz, et al., 2018; Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b) and likely linked to the underestimation by DGVMs of the land sink enhancement in response to the eruption of Mt. Pinatubo (Lucht et al., 2002; Mercado et al., 2009). Other years are associated with positive and negative phases of El Niño (Bastos et al., 2018; Bowman et al., 2017), suggesting a possible contribution to D_{Globe} of mismatches in the response of DGVMs and inversions to ENSO.

The "budget imbalance" time series from GCB2018 (black line, Figure 1) shows very similar variability to the group of ensemble means, supporting the hypothesis that errors in the land sink representation explain a large fraction of the GCB2018 imbalance term. The values of *D* show an offset compared to this term though, which is in part explained by the adjustment of lateral fluxes to Resplandy et al. (2018) values. Even though over the 39-year period no clear trend can be distinguished in the whole ensemble, seven out of the 80 pairs inversion-DGVM show significant trends, evaluated by the Mann-Kendall test.

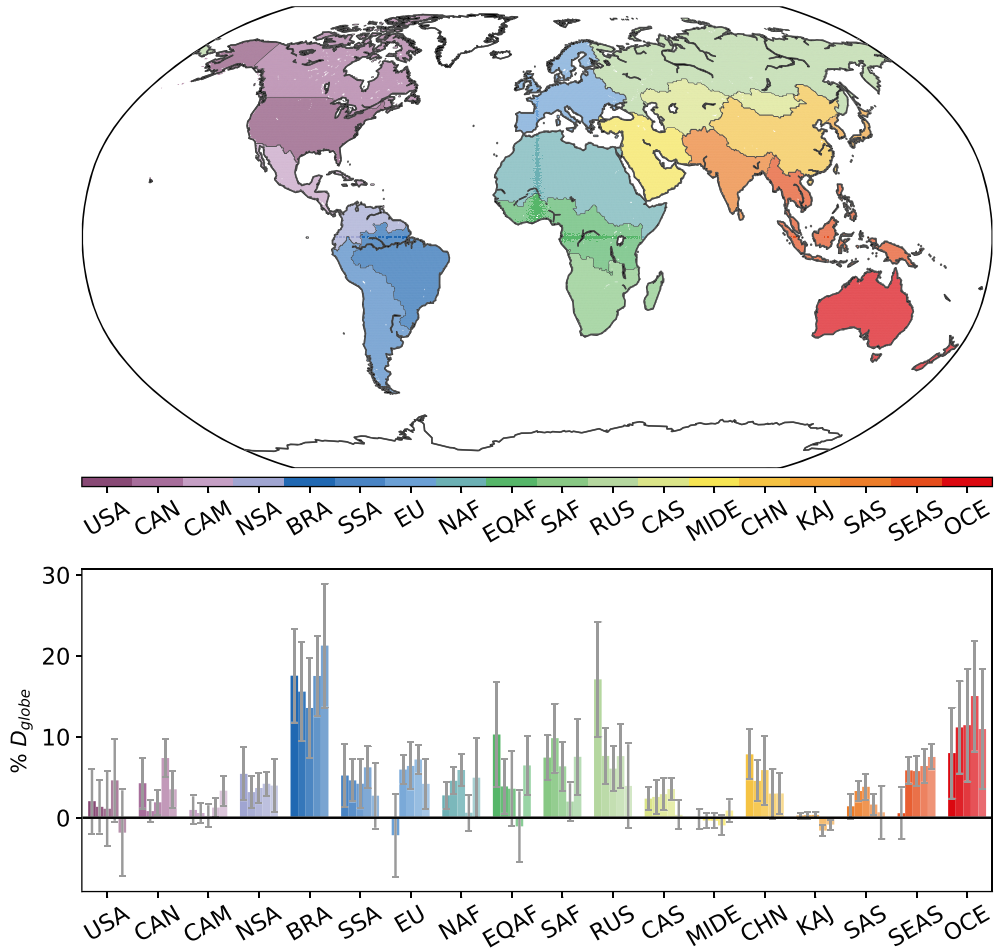

Figure 2. The 18 study regions: United States (USA), Canada (Canada), Europe (EU), Northern Africa (NAF), central Asia (CAS), Russia (RUS), Korea and Japan (KAJ), China (CHN), Southeast Asia (SEAS), Oceania (OCE), south Asia (SAS), Middle East (MIDE), southern Africa (SAF), equatorial Africa (EQAF), southern South America (SSA), Brazil (BRA), northern South America (NSA), and central America (CAM) are shown in the top panel. In the lower panel, the contribution of each region to the variance of D is shown for each inversion, calculated for the multi-DGVM mean for each of the five inversions. The bar order corresponds to CAMS, CarboScope s76 and s85, MIROC, and CarbonTracker Europe.

The variability across (inversion, DGVM) pairs is much higher than the variability from year-to-year of the ensemble means. The 1σ of D_{Globe} is 0.9 Pg C yr^{-1} (Figure 1). This value is comparable to the uncertainty reported for the GCB2018 imbalance term ($\pm$1 Pg C yr^{-1}).

3.2. Regional Contributions to the Global Differences

We quantify the contribution of each of the 18 regions to the variance of the spatially averaged differences over the globe (D_{Globe}, Figure 2) for the $n_{inversions} \times 16$ DGVM pairs. All (inversion, DGVM) pairs single out Brazil (BRA) as the region contributing the most to D at the global scale (16–27%, averaged accross DGVMs), followed by Oceania (OCE, 7–17%). Inversions also agree on a moderate contributions of northern South America (NSA, 4–7%), southern South America (SSA, 3–6%) and China (CHN, 3–11%). Four out of five inversions estimate a strong contribution of Southeast Asia (SEAS) to D_{Globe} (−1% for CAMS, 6–11% for the other four). CAMS assigns a strong negative contribution of EU (−6%) versus a strong positive contribution of RUS (21%) to D_{Globe}, while the other inversions indicate comparable contributions of EU and RUS (0–7%). This suggests that inversions have dipoles in their allocation of the Eurasian CO_2 sink with inversions having more uptake in Russia and less in Europe. This is likely due to the relatively sparse observation network in RUS that does not allow to separate easily the flux of nearby regions.

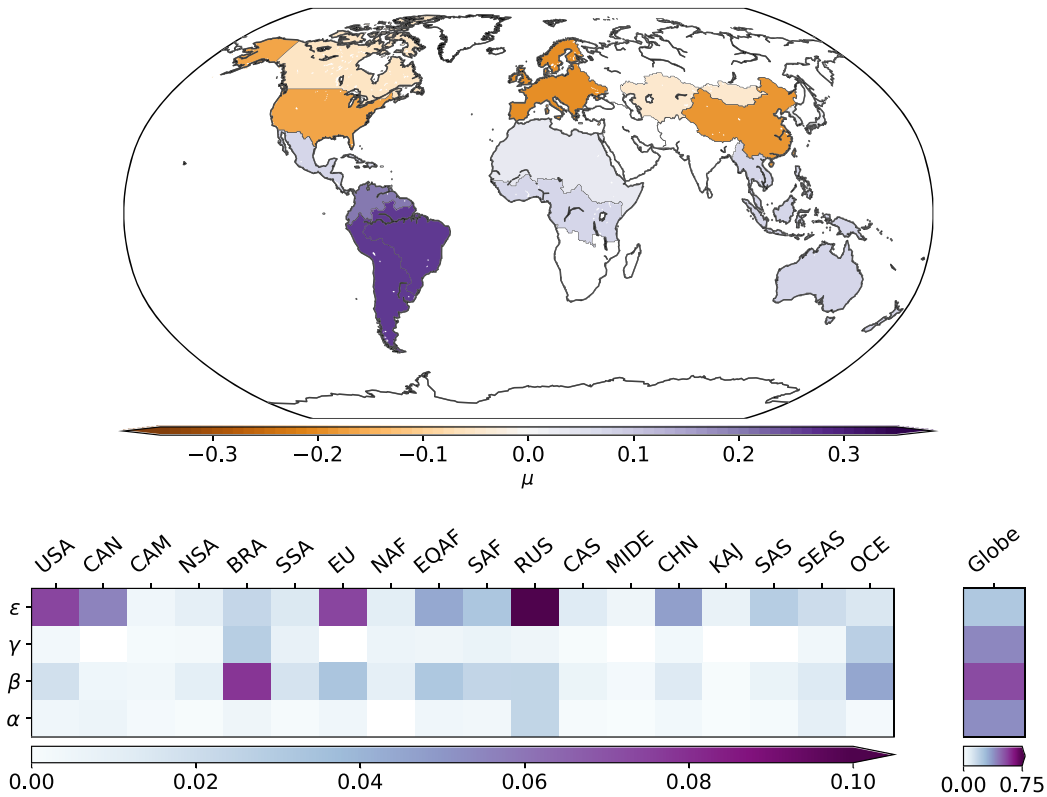

Figure 3. Coefficients of the linear mixed effects model fit only with random effects only (equation (4)) to the time series of D over each region. The intercept μ corresponds to the error D averaged over all models and all years over each region (top panel). A positive (negative) value of μ indicates an overestimate (underestimate) of NBP by the DGVMs compared to inversions over the study period (1979–2017). Regions where μ values are not significantly different from zero are masked out in white. The coefficients for the random effects for each region and for the globe are shown in the bottom panel. A high value indicates a strong contribution of random effects from differences between DGVMs (α), between-year variability (β), model year variability (γ), and variability between inversions (ϵ).

If the contribution of a given region to the variance of D_{Globe} changed over time, differences between the three "long" and the two short inversions (Table 1) can be explained partly by their different validity periods. However, large differences are found even between inversions with comparable validity periods, so they are most likely due to differences in the inversion systems. Moreover, we find a large spread of results between (inversion, DGVMs pairs) for several regions, indicating a relevant contribution of the uncertainty in spatial patterns of NBP from DGVMs to D_{Globe}.

As for the globe, the spread of regional D (D_{reg}) for the 5×16 members ensemble in certain regions can be very high (Figure S2). There is not necessarily a correspondence between the regions with larger D_{reg} spread and those regions identified above for their contribution to D_{Globe}. The regions showing larger average spread over the 39-year period are, in decreasing order, RUS (1.3 Pg C yr^{-1}), USA, and EU (1.0 Pg C yr^{-1}), and BRA and CAN (0.9 Pg C yr^{-1}). In some years the spread in D_{reg} can reach 2.6 and 1.8 Pg C yr^{-1} for RUS and BRA, respectively, and 1.6 Pg C yr^{-1} in USA, EU, and CHN, which is close to the magnitude of the global sink (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b). In most regions, it is hard to distinguish between individual inversions, showing that the variability across DGVMs also contributes considerably to the range of D_{reg}. Moreover, some regions show consistent interannual or long-term variability in D_{reg} and its spread. For example, D_{BRA} is predominantly positive in the 1980s and 1990s, and becomes mostly negative following 2010. In OCE, a very large spread across data sets is observed for some years (e.g., 1999 and 2011) but not during most of the study period.

3.3. Decomposition of D

The intercept (μ) of the LME_{RE} fit corresponds to the average stationary value of D_{reg} for the period 1979–2017 (Figure 3 top panel). For the globe, μ is positive, indicating that DGVMs estimate a stronger global CO_2 sink than inversions. Regionally, negative μ values are found over most of the Northern Hemisphere's

extratropics (except CAN and RUS) and positive values in the tropics and southern extratropics. Negative (positive) values indicate that DGVMs report a weaker (stronger) sink in the northern (tropical and southern) regions. The largest μ terms are found in South America (> 0.2 Pg C yr^{-1}), followed by EUR, CHN and USA (< −0.15 Pg C yr^{-1}). The tropics and southern extratropics have average D of 1.0 Pg C yr^{-1} and the northern regions of −0.5 Pg C yr^{-1}.

The variance of D_{Globe} and D_{reg} attributed to each random term of $\mathrm{LME_{RE}}$ is shown in Figure 3 (bottom panel). Interannual variability (β) is the term contributing the most to D_{Globe}, followed by the between-DGVM variability (α) and model year combined (γ). Variability between inversions (ϵ) contributes the least to D_{Globe}, but at the regional scale it is the most important factor in half of the regions, and is especially high in USA, CAN, EU, RUS, and CHN.

Interannual variability is the second most relevant term regionally, especially in BRA, OCE, NSA, and SSA where the interannual term dominates over the difference between inversions, although for the latter two the variance is rather small. Between-DGVM differences do not contribute much to the regional differences, except in RUS.

If random effects could explain all of the variance of D, the residuals of the model fit should be randomly distributed around 0 and not change over time. However residuals of the $\mathrm{LME_{RE}}$ fit show clear trends or multiannual patterns (Figure S3). This misfit of $\mathrm{LME_{RE}}$ suggests that deterministic processes that explain interannual variability in D are missing or have a common bias in either inversions or DGVMs (probably being attributed to β). The most evident cases are the decadal variations in the residuals for the globe, the sharp decreasing trend in the misfit in BRA since circa 2005, and long-term trends in SSA, EU, and RUS.

3.4. Effects of Predictors

For each region and for the globe, we test a set of key predictors (fixed effects model, $\mathrm{LME_{FE}}$, equation (5)) that might help explain D. The predictors tested are the (i) a simple linear trend ("year" as predictor); (ii) the Oceanic Nino Index (ONI); (iii) F_{LUC} from the BLUE bookkeping model between 1979 and 2017, and (iv) the same predictors plus E_{fire} for 1997–2017 (the period covered by GFED4.1s). For each region, we tested additional predictors such as TWS, burned area and changes in forest and cropland areas. These additional predictors are correlated with the previous ones (e.g., soil water with ENSO, or F_{LUC} from cropland/forest area changes) and were not found to provide additional information. The regression coefficients of the $\mathrm{LME_{FE}}$ fit to D_{reg} and D_{Globe} are shown in Figure 4 for the two periods.

The coefficients indicate the effect of a unit change in the predictor (e.g., F_{LUC}) to a corresponding unit change in D. We find that D_{Globe} is dominated by ONI and F_{LUC}, with a small trend component, during 1979–2017, and only by the trend during 1997–2017, as other effects cancel out regionally. All coefficients for the globe have negative sign, indicating that positive anomalies of each predictor contribute to explain an underestimate of the CO_2 sink by DGVMs relative to inversions. This means, for example, that a positive anomaly in land use change emissions (above average) would coincide with lower NBP in DGVMs compared to inversions.

For the 39-year period, a strong negative relationship is found between D_{Globe} and ONI, implying that DGVMs tend to estimate a weaker sink or stronger source than inversions in response to El Niño, and the opposite during La Niña events. The global relationship of D_{Globe} and ONI is mainly driven by NSA, BRA, SAF, and SEAS, which also show strong negative effects of ONI, and partly offset by an opposing effect in CAS, MIDE, and SAS.

The LME_{FE} indicates D_{Globe} is explained by emissions from land use change in those regions where land use changes are more intense. The minus sign indicates that the higher F_{LUC} (positive for emissions), the more DGVMs underestimate the net sink globally compared to inversions. Such a negative contribution from LUC emissions is found in BRA, EU, EQAF, and SEAS. We tested $\mathrm{LME_{FE}}$ using F_{LUC} simulated by DGVMs (Experiments S2 and S3), which yielded very similar results. This suggests that too strong emissions from LUC in DGVMs might result in a weaker sink compared to inversions (which implicitly include F_{LUC}).

We find a weak negative trend contribution to D_{Globel}, revealing a significant, though small, divergence of inversions and DGVMs over time by −26 Tg C decade^{-1}. In other words, DGVMs underestimate the rate of increase of the global terrestrial sink compared to inversions. Significant trend components of D are found in 8 regions (negative in NSA, BRA, SSA, RUS, and CAS and positive in EU, KAJ, and SEAS), but their values are rather small (from 1–10 Tg C yr^{-1} in absolute magnitude).

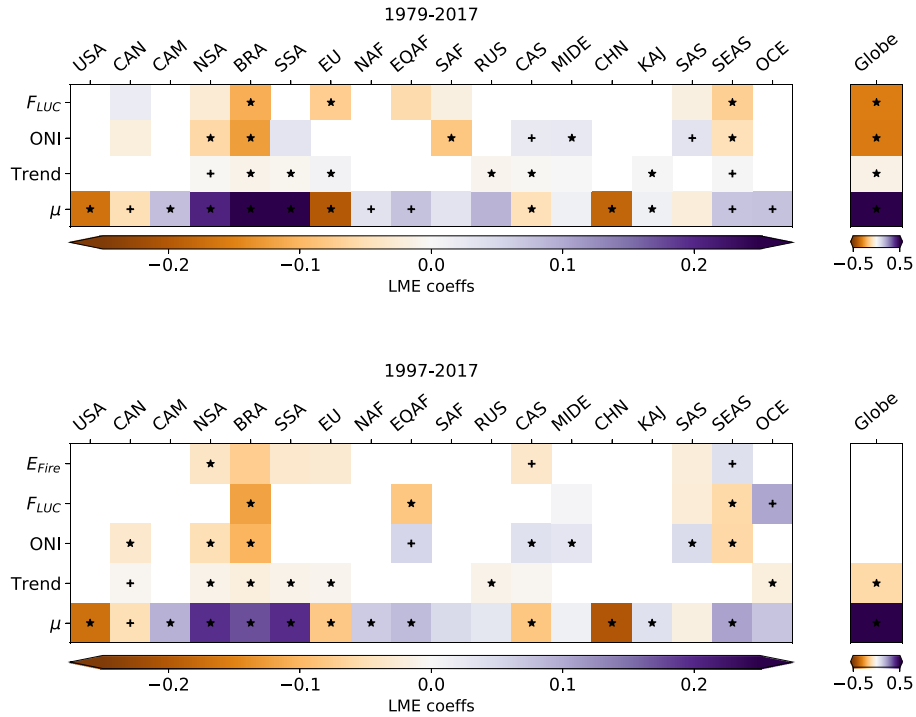

Figure 4. Coefficients of the LME model fit with fixed effects (equation (5)). F_{LUC}, ONI, and E_{Fire} correspond to emissions from land use change, the Oceanic Niño Index, and emissions from fire, respectively. The regions abbreviations are described in Figure 2 . The symbols indicate statistically significant results: crosses for p value < 0.05 and asterisks for p value < 0.01. The top panel shows values for the period 1979–2017, and the bottom panel for the period 1997–2017.

Significant effects of E_{fire} are found in NSA, CAS, and SEAS (Figure 4), the first two negative (DGVMs estimating lower sink/stronger source associated with fires than inversions) and positive in SEAS (DGVMs estimating stronger sink/weaker source associated with fires than inversions). Including E_{fire} reduces the variance of the residuals of the LME fit to D_{Globe} during 1997–2017 by 80%, compared to a reduction of 58% for the fit without fire in that period (Figure S3). For one DGVM simulating fire using a state-of-the-art fire module, Yue et al. (2015) have shown a tendency of the model to overestimate fire emissions and burned area in regions roughly corresponding to NSA, BRA, and CAS, and to underestimate E_{fire} and burned area in SEAS. This is discussed in section 3.5.3.

3.5. Sources of Uncertainty

3.5.1. ENSO, Temperature, and Soil Moisture Variability

The sensitivity of soil moisture from DGVMs to ONI in the regions with significant effects (NSA, BRA, SAF, and SEAS, negative and CAS, MIDE, and SAS, positive) is generally consistent with that of TWS from the GRACE reconstruction (Humphrey et al., 2018), although slightly underestimated for NSA and BRA (Figure S4). In SAF and SEAS the sensitivity of SM from DGVMs to ONI is close to that of TWS, but in the former DGVMs show a large range. The differences between DGVMs and inversions in these regions can be further related with the sensitivity of modeled NBP to the hot/cooler (e.g., by too strong respiration response to warming) or dry/wet anomalies linked with El Niño/La Niña events (e.g., by too strong water stress controls on productivity). Therefore, we evaluate the sensitivity of NBP to temperature and simulated soil moisture.

In NSA and BRA, inversions tend to estimate a positive sensitivity of NBP to temperature (higher uptake for warmer conditions, Figure 5), indicating a positive effect of temperature on productivity, which is mainly limited by radiation rather than water (Nemani et al., 2003). On the contrary, DGVMs tend to estimate negative sensitivity of NBP to temperature (lower NBP with warming). On the other hand, DGVMs estimate higher sensitivity of NBP to water availability than inversions (Figure 5). Higher sensitivity to water availability is also observed in SEAS, but inversions and DGVMs show similar sensitivity to temperature in this region.

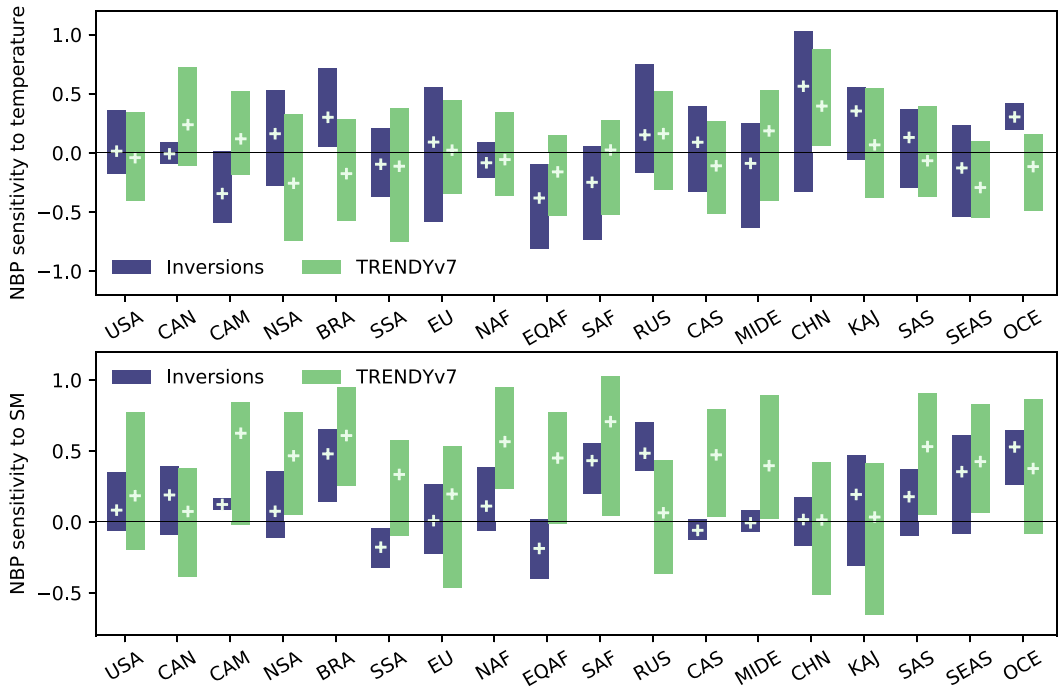

Figure 5. Sensitivity of NBP to annual temperature (CRU-JRA, top panel) and to soil moisture for inversions (GRACE reconstruction from (Humphrey et al., 2018)) and DGVMs (with simulated soil moisture) (bottom panel). The bars indicate the range of sensitivities from each set of data, and the crosses show the ensemble mean.

3.5.2. Land Use Change

The results of LME_{FE} (Figure 4) point to a negative effect of F_{LUC} in D_{Globe}, indicating that DGVMs might overestimate emissions from LUC or underestimate LUC-related recovery sinks. However, the F_{LUC} data used relies on the same LUC forcing as DGVMs (BLUE) as a predictor. Therefore, the agreement between BLUE and DGVMs' estimates of F_{LUC} is partly explained by their use of a common forcing. In this sense, BLUE cannot be seen as a fully independent estimate of F_{LUC}. We find a remarkable disagreement between regional estimates of F_{LUC} from DGVMs and BLUE with those from HN2017, especially in NSA, BRA, and SEAS. This points to a strong influence of the underlying LUC data used to force models, more than to structural differences between bookkeeping models or differences between bookkeeping and process-based models.

We compare F_{LUC} from DGVMs with the two bookkeeping models (BLUE and HN2017, $F_{LUC-BLUE}$ and $F_{LUC-HN2017}$, Figure 6). Simulation S3 (climate, CO_2 and LUC), although more suitable to compare with NBP from inversions, includes in their F_{LUC} the LASC term ($F_{LUC-LASC}$). The difference between S3 and S2 (climate and CO_2 only) is thus not directly comparable to bookkeeping model estimates. Therefore, we further include results from Simulation S4 (LUC only under preindustrial climate and CO_2) that does not include the LASC to calculate $F_{LUC-noLASC}$.

Globally, the multimodel ensemble mean (MMEM) of $F_{LUC-LASC}$ is above the estimates by the two bookkeeping models during most of the 39-year period (Figure 6), but $F_{LUC-noLASC}$ values are mostly between the two bookkeeping models, when they would be expected to be slightly lower. The two bookkeeping models differ on average by 0.5 Pg C yr^{-1}, with BLUE systematically above HN2017, possibly because it includes more processes, particularly shifting cultivation (Arneth et al., 2017), or because HN2017 use grasslands preferentially over forest for pasture expansion (Hansis et al., 2015).

In BRA, the LASC has a relatively small effect (Figure 6). The MMEM for $F_{LUC-noLASC}$ is generally close to the estimates of BLUE and shows the same dynamics: a decreasing trend from the 1980s until circa 2009, and then a sudden increase in F_{LUC} of about 0.5 Pg C yr^{-1}, followed by a period with high emissions and strong interannual variability. This 2009–2017 period coincides with a strong decrease in D_{reg} in BRA (Figure S2),

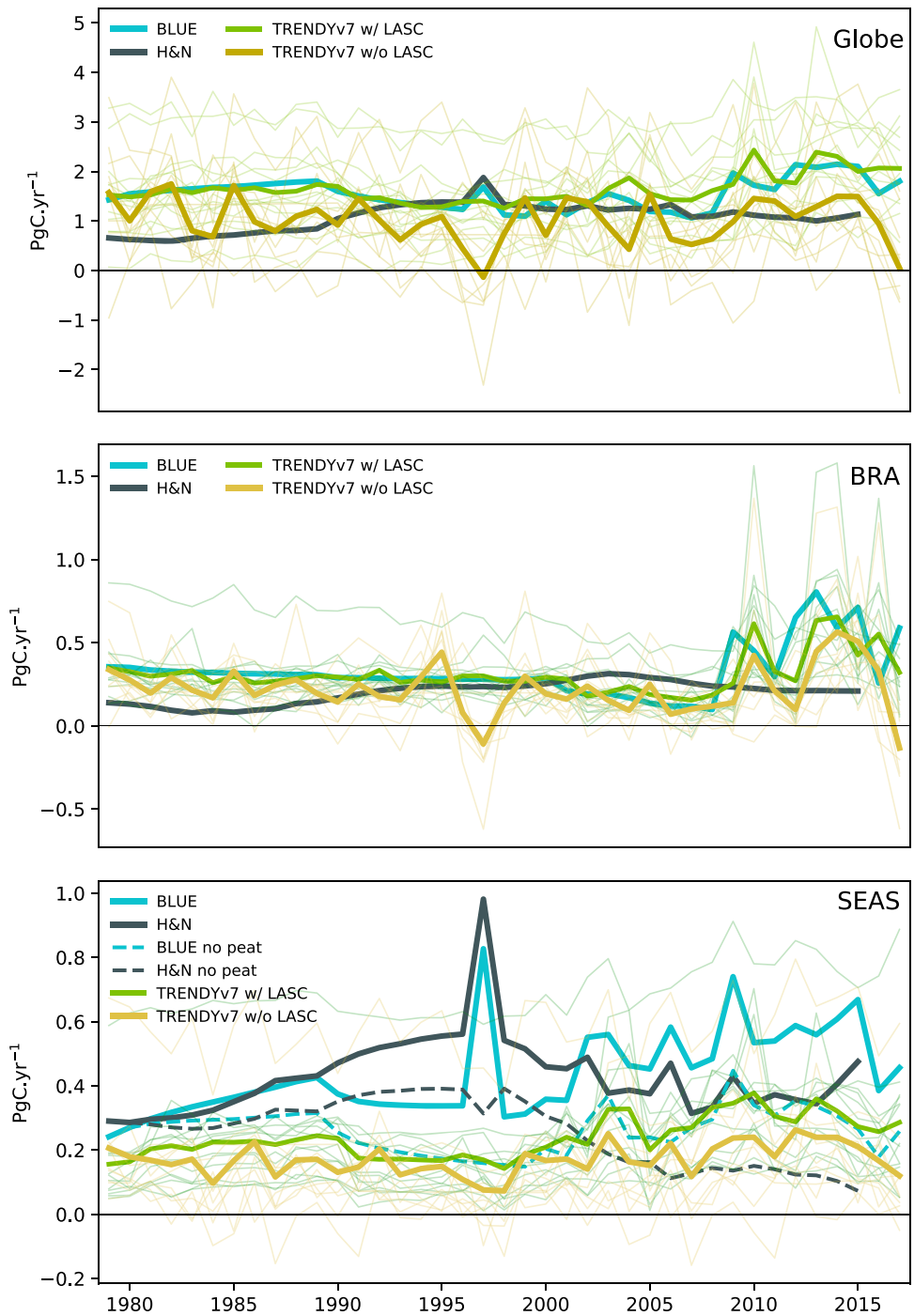

Figure 6. Comparison of F_{LUC} from different data sets for the (top) globe, (middle) BRA, and (bottom) SEAS (where F_{LUC} is a significant predictor of D).

that is, DGVMs simulating a much weaker sink than inversions. HN2017, on the contrary, indicates lower emissions in the 1980s, a slow increasing trend until circa 2005 and a decreasing trend afterward.

In SEAS, peat burning and drainage are a major component of LUC-related fluxes (Moore et al., 2013). Since DGVMs do not simulate peat, the comparison of F_{LUC} with bookkeeping models should be adjusted for peat burning and drainage emissions (also note that part of these fluxes in models could be included in E_{fire}, which are discussed in the next section). When these emissions are removed from F_{LUC} estimated by bookkeeping models (Figure 6): $F_{LUC-LASC}$ is close to the higher value estimated by BLUE, and $F_{LUC-noLASC}$

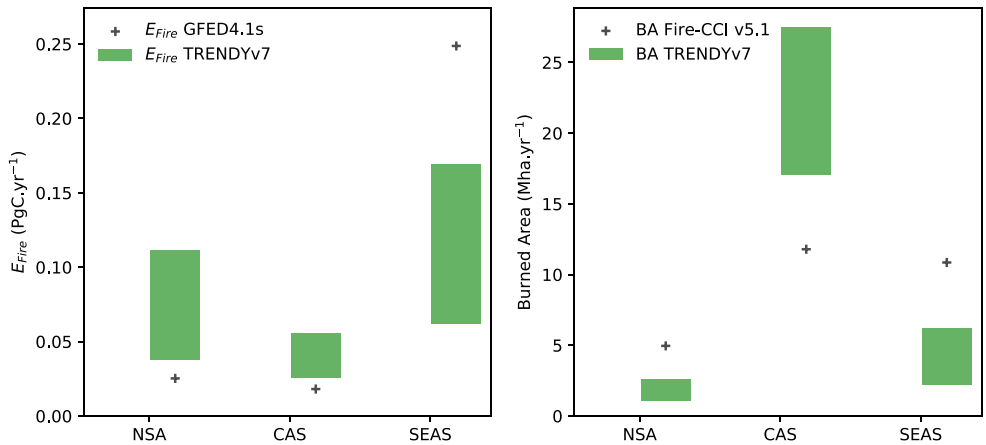

Figure 7. Comparison for the three regions where fire has a significant effect on D of E_{fire} (cross markers, left panel) and burned area (cross markers, right panel) simulated by DGVMs (green bars indicate the DGVM range) and E_{fire} reported by GFED4.1s and burned area from Fire-CCI Burned Area v5.1.

is between the two bookkeeping models, but only following the year 2000. Before that, $F_{LUC-noLASC}$ is well below both bookkeeping models, which is consistent with the lower equilibrium C densities in this simulation, as compared with bookkeeping values. The year 2000 also marks a period with higher F_{LUC} in BLUE, indicating a possible effect of the LUC forcing.

3.5.3. Fire

Fire emissions from GFED4.1s are found to be a significant predictor for D_{reg} in NSA, CAS and SEAS. In the first two regions, E_{fire} effects on D_{reg} are negative, indicating that higher fire emissions in GFED4.1s are associated with too weak CO_2 uptake by DGVMs compared with inversions. In SEAS, the effect is the opposite, which suggests an underestimation by DGVMs of E_{fire}.

In Figure 7 (left panel) we compare E_{fire} from those DGVMs that simulate fire (7 out of 16) with GFED4.1s fire emissions (Giglio et al., 2016). In NSA and CAS the ensemble of DGVMs estimates fire emissions about twice as high on average as those reported by GFED4.1s, but also much more variable (Figure S6). On the contrary, in SEAS, DGVMs estimate fire emissions that are, on average lower than the total emissions reported by GFED4.1s (about 60%). These results are consistent with overestimation of fire emissions in NSA and CAS leading to an underestimation of the CO_2 sink compared with inversions, and the opposite in SEAS.

Biases in emissions from fire can be due to biases in simulated burned area or in the fuel properties and fire characteristics (e.g., human influence on ignitions). The comparison of simulated burned area with Fire-CCI burned area product (Chuvieco et al., 2016) indicates an overestimation by DGVMs of burned area in CAS and underestimation in SEAS, consistent with E_{fire} biases. However, this is not the case in NSA, where DGVMs estimate too high fire emissions in spite of underestimating burned area.

4. Discussion

Tropical regions explained most of the variance in D_{Globe}, especially Brazil and northern South America, Southeast Asia and Oceania. These are, at the same time, the regions less constrained by observations. However, the regions with higher biases (Figure 3) are generally those with better observational coverage (USA, CAN, EU). These regions are strong E_{FF} emitters, suggesting that spatial and temporal differences between prescribed E_{FF} might still affect surface fluxes estimated by inversions, which were only adjusted for latitudinal bands.

At regional scale, negative values of D (DGVMs underestimating the long-term sink) dominate in the Northern Hemisphere (especially Europe, China and USA), and positive values in the south, mostly in South America. We find that regional differences can be mostly attributed to variability across inversions, followed by interannual variability in the fluxes.

We found a strong contribution of interannual variability in NBP to D, associated with El Niño–Southern Oscillation (ENSO) and fluxes from land use change (F_{LUC}), especially in Brazil and Southeast Asia, the two regions also contributing the most to D_{Globe} variance.

Using a similar set of DGVMs and some of the inversions used here, Bastos et al. (2018) have shown that DGVMs estimated a positive source anomaly in response to El Niño in 2015/2016 that was 0.3–0.6 Pg C yr^{-1} stronger than inversions. The coefficient of ONI estimated here (−0.3 Pg C yr^{-1}/s.d.u., s.d.u. being the ONI values in standard deviation units) would imply a source anomaly difference between the inversions and DGVMs of about 0.5 Pg C yr^{-1} for these years ($ONI_{2015/16} = 1.6 s.d.u.$), consistent with their estimate.

Several studies have shown that variability in tropical CO_2 fluxes (mainly driven by ENSO) is better explained by variations in soil water storage than by temperature alone (Bastos et al., 2013; Humphrey et al., 2018; Poulter et al., 2014). The regions with stronger negative coefficients for ONI (NSA, BRA, SAF, and SEAS, Figure 4) are those where El Niño events impose strong warm and dry conditions, and those with positive coefficients (CAS, MIDE, and SAS) are associated with cool and wet conditions during El Niño (Bastos et al., 2013; Mason & Goddard, 2001). The combination of these two results hints at differences in the sensitivity of NBP from inversions and DGVMs to changes in water availability or temperature related with ENSO (section 3.5.1), or possibly also to changes in atmospheric circulation/mixing patterns linked to ENSO. The opposing sign in the sensitivity to temperature between inversions and DGVMs (Figure 5) may be due to an overestimate of the sensitivity of respiration to temperature (Bastos et al., 2018; van Schaik et al., 2018), or indirectly linked to the higher sensitivity of NBP to water availability in DGVMs compared to inversions. Models differ in prescribed soil depth and root water access for transpiration, but in general do not have deep rooting and ground water access of plants (Fan et al., 2017), which may explain why the sensitivity of regional CO_2 fluxes to water availability in DGVMs is higher than that of inversions with TWS. Still, inversions also show high uncertainty in these regions, which hampers the attribution of errors to one or other data source.

Excluding the years corresponding to the Mt. Pinatubo eruption, the strong peaks in D_{Globe} are associated with moderate or strong El Niño (1982, 2009/2010, and 2015) or La Niña events (1984, 1988/1989, 2000, and 2011). In most of these years,the GCB2018 budget imbalance term shows similar changes as D_{Globe}. This suggests that improvements in DGVMs modeling of tropical vegetation sensitivity to water availability could help reducing this gap in future GCBs.

The higher values of $F_{LUC-LASC}$ are consistent with the effect of CO_2 fertilization on natural ecosystems resulting in higher LUC-related emissions if a pristine forest is converted to managed land. However, $F_{LUC-noLASC}$ values (Figure 6) remain mostly between the two bookkeeping models, when they would be expected to be slightly lower, given that they rely on preindustrial C densities. Since most DGVMs do not include shifting cultivation, which is included in the BLUE model, estimates of F_{LUC} from DGVMs are likely too high. The negative effect of F_{LUC} to D_{reg} in BRA and SEAS, and consequently in D_{Globe}, suggests therefore a possible overestimation of LUC-related emissions due to the forcing used, for example, higher deforestation rates in LUH2 than in FAO/FRA.

The strong peak in emissions from LUC following 2009 in DGVMs and in BLUE is mainly explained by sharp increases in cropland area during 2009–2017 coinciding with increased forest loss in the LUH2v2.1h land cover data set (Figure S5). On the contrary, in the FAOSTAT (2015) data used by HN2017, most of the cropland area increase occurs before 2009, and forest loss slows down only after 2005. The LC-CCI also reports a slow down of forest area loss and stabilization of cropland in Brazil after 2005. A slowdown in deforestation rates in Brazil over this period has further been reported by Hansen et al. (2013) using high-resolution satellite imagery. This indicates that the overestimation of DGVMs of F_{LUC} inferred from the MMEM appears to be related with biases in deforestation rates from LUH2v2.1 in this region.

In SEAS, the two forcings show opposing changes in deforestation rates starting from circa 2000 onward (Figure S5), coinciding with higher F_{LUC} in DGVMs and BLUE (Figure 6) LUH2v2.1h indicates a faster decrease of forest area, while (FAO, 2015) shows a tendency for deceleration in deforestation. LUH2v2.1h also reports higher cropland expansion values, with much faster rates following 2005. The study by Hansen et al. (2013) indicated that gross forest loss rates increased in Indonesia between 2000–2012. Li et al. (2018) showed that net forest area changes in Southeast Asia and Indonesia from Houghton et al. (2012) and from LUH2v2h between 1992–2012 are consistent with the changes reported by Hansen et al. (2013), but much

stronger than reported in LC-CCI data. LUH2v2h also reports higher increase in cropland areas when compared to LC-CCI (Figure S5). Li et al. (2018) pointed that these differences could be from the increase in plantations such as oil palm, which is categorized as cropland in FAOSTAT but detected as forest from satellites. It is thus not fully clear if the higher emissions by DGVMs in SEAS compared to HN2017 are indeed because of the LUC forcing.

We found that biases in burned area explain well biases in E_{fire} and its contribution to D_{SEAS} and D_{CAS}, but not in NSA. A possible explanation for this disagreement might be related with simulated soil moisture and fuel dryness in this region (Van Leeuwen et al., 2013). For example, some models show very high E_{fire} in 1997/1998 and 2015 in NSA compared to GFED4.1s. These very strong peaks are possibly a consequence of how vegetation response to drought during El Niño and consequent fire emissions (Patra et al., 2005) are simulated by DGVMs, perhaps leading to an overestimation of fire intensity and combustion completeness (since burned area is still lower than that of Fire-CCI in those years). In order to understand the errors in simulated fire emissions by DGVMs, a much more detailed analysis of the different factors controlling E_{fire}, from burned area to fuel characteristics and fire dynamics, is needed, which is beyond the scope of this study.

DGVMs tend to simulate a slower increase in CO_2 uptake than reported by inversions in several regions and globally. The weak negative trend in D_{Globe} is consistent with studies showing that DGVMs may underestimate the response of vegetation to the effect of CO_2 fertilization (Fernndez-Martnez et al., 2019; Thomas et al., 2016), or the effect of N deposition in Asia (Liu et al., 2013), or forest regrowth (Pugh et al., 2019). The positive values for the trend in EU, KAJ, and SEAS indicate that DGVMs show a stronger long-term increase in the land sink compared to inversions. This might be because DGVMs underestimate the negative effects of warming and decrease in water availability concurrent with increasing CO_2 in some regions (Buermann et al., 2018). However, inversions might also indirectly overestimate the land sink by underestimating the ocean sink intensification in recent years (Landschützer et al., 2015).

We found a significant contribution of uncertainty in F_{LUC} and of E_{fire}, as well as a small trend component, to regional and global D. It is thus worth evaluating whether those DGVMs including representation of key LUC processes such as shifting cultivation and wood harvest (Arneth et al., 2017), N deposition or fires perform better compared to the ensemble of 16 DGVMs.

By fitting the LME to each subset of DGVMs, we find that simulating fire, N deposition or wood harvest does not lead to major reductions in uncertainty (Table S1). The DGVMs including shifting cultivation, though, show significant reduction of the bias globally, as well as in SEAS and EQAF, regions where these practices are relevant (Heinimann et al., 2017). When adding the process representation as a predictor in the LME_{FE} (Figure S7), including some of these processes as predictors are found to improve the LME fit, but the coefficients are generally nonsignificant, excepting N deposition globally and wood harvest in some regions. These results indicate that including gross transitions in F_{LUC} estimates is a key improvement to DGVMs, in order to reduce uncertainties in global and regional budgets.

The fact that models with fire do not perform much better than those without fire may be because models do not realistically simulate burned area and E_{fire}, as discussed above, but also because those DGVMs not simulating fire might compensate by having higher sensitivity of decomposition to temperature (as found for tropical regions in Figure 5). Simulating wood harvest also does not seem to necessarily reduce the bias in D, and is only a weak predictor in NSA, NAF and KAJ. This is possibly because the wood harvest fluxes are generally small, compared to other terms, for example, shifting cultivation (Wilkenskjeld et al., 2014), or because trade of wood products is not accounted for, and therefore the location of sinks and sources related with wood production and consumption is not well captured by DGVMs. Finally, N deposition if found to significantly contribute to D_{Globe} (the negative sign indicating that models simulating N deposition tend to have lower D), but when fitting the LME_{FE} model to that subset of DGVMs, no reduction in D is found. This suggests that uncertainty from other terms is probably higher.

5. Conclusions

In this study we attempted to identify the sources of uncertainty explaining the budget imbalance term of the latest GCB (Le Quéré, Andrew, Friedlingstein, Sitch, Hauck, et al., 2018b). We compared DGVM outputs of the NBP between 1979 and 2017 with results from an ensemble of atmospheric inversions, whose fluxes are consistent with the growth rate of atmospheric CO_2. We showed that the difference between NBP estimates

by DGVMs and inversions matches the main variation patterns of the GCB2018 budget imbalance term. At the global scale, interannual variability and between-DGVM variability contribute the most to *D*, and variability across inversions contributes the least. Based on our results, we could identify key regions and processes that deserve special attention from the Earth System and C cycle communities:

(i) the reduction of uncertainty in atmospheric inversions, for example, through more in situ observations in the tropics, or the use of remote-sensing;
(ii) improvements in F_{LUC} estimates, by including shifting cultivation DGVMs and also by evaluating thoroughly land use change estimates at country level, especially for the period overlapping high-resolution satellite records;
(iii) the improvement of tropical hydroecological processes and fire representation within DGVMs;
(iv) accurate estimates of lateral carbon fluxes and their variability, to guarantee a fair comparison between top-down and bottom-up approaches.

There are several ongoing efforts to tackle each of the processes highlighted above. Satellite-based inversions might help to reduce uncertainty in the flux partitioning between land, ocean, and different continents, and regional inversions might provide better constraints of regional CO_2 fluxes. Above-ground biomass products at annual time steps and moderate resolution are currently available at least for the tropics (Brandt et al., 2018), providing a truly independent estimate of losses and gains of carbon in ecosystems due to both natural and anthropogenic processes. Considerable work is being put into improving tropical phenology, vegetation-water interactions, mortality, and nutrient limitations by several modeling teams.

The predictors tested in this work reduced the variance of residuals of *D* only by 25–51%, and residuals still show large variability at the global and regional scale. This points to other processes not discussed in this study either on the DGVM side (e.g., response of forests to climate change, forest regrowth or forest aging, or their response to nitrogen deposition to nutrient limitations) or on the inversion side (e.g., errors in transport modeling, inversion set up, atmospheric monitoring network) that might deserve more attention in the future. Such efforts will hopefully allow reducing uncertainty in global and regional CO_2 budgets in the near future. This, in turn, is fundamental for a science-based evaluation of the effectiveness of mitigation policies implemented in light of the Paris Agreement.

Acknowledgments
This work was partly supported by the European Space Agency Climate Change Initiative ESA-CCI RECCAP2 project (ESRIN/ 4000123002/18/I-NB). P. K. P. is partly supported by the Environment Research and Technology Development Fund (2-1701) of the Ministry of the Environment, Japan. A. J.'s effort was supported in part by the Department of Energy (DESC0016323) and NSF (NSF AGS 12-43071). P. C. acknowledges support from the European Research Council Synergy Project SyG-2013-610028 IMBALANCE-P, and P. C. and D. M. the ANR CLAND Convergence Institute. W. P. received funding from the European Research Council (ERC) for the Airborne Stable Isotopes of Carbon from the Amazon (ASICA) project, Contract 649087. N. E. S. received funding from the Netherlands Organisation for Scientific Research (NWO) for the Ruisdael Observatory. I. T. L. received funding from NWO under Contract 016.Veni.171.095. CT Europe simulations were performed using a grant for computing time (SH-312, 16666) from NWO. The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. Oak Ridge National Laboratory is operated by UT-Battelle, LLC, under Contract DE-AC05-00OR22725 with the U.S. Department of Energy. The authors would like to thank Corinne Le Quéré and Johannes Winckler for constructive comments in the preparation of this study and the ESA Soil Moisture CCI and the Fire-CCI teams for producing and maintaining the data sets used. The GRACE reconstructed data set was kindly provided by V. Humphrey.

References

Andres, R. J., Boden, T. A., Bron, F.-M., Ciais, P., Davis, S., Erickson, D., et al. (2012). A synthesis of carbon dioxide emissions from fossil-fuel combustion. *Biogeosciences*, *9*(5), 1845–1871. http://www.biogeosciences.net/9/1845/2012/

Arneth, A., Sitch, S., Pongratz, J., Stocker, B., Ciais, P., Poulter, B., et al. (2017). Historical carbon dioxide emissions caused by land-use changes are possibly larger than assumed. *Nature Geoscience*, *10*(2), 79.

Bastos, A., Sullivan, M., Ciais, P., Makowski, D., Sitch, S., Friedlingstein, P., Chevalier, F., Rdenbeck, C., Pongratz, J., Luijkx, I., Patra, P., Peylin, P., Canadell, J., Lauerwald, R., Li, W., Smith, N., Peters, W., Goll, D., Jain, A., Kato, E., Lienert, S., Lombardozzi, D., Haverd, V., Nabel, J., Tian, H., Walker, A., & Zaehle, S. (2019). Aggregated regional estimates of net atmosphere-land CO2 fluxes from the five atmospheric inversions and 16 Dynamic Global Vegetation Models, supplemental data to Bastos et al, 2019 (https://doi.org/10.1029/2019GB006393). https://doi.org/10.18160/1SVH-3DNB

Bastos, A., Friedlingstein, P., Sitch, S., Chen, C., Mialon, A., Wigneron, J.-P., et al. (2018). Impact of the 2015/2016 El Niño on the terrestrial carbon cycle constrained by bottom-up and top-down approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20170304.

Bastos, A., Running, S. W., Gouveia, C. A., & Trigo, R. M. (2013). The global NPP dependence on ENSO: La Niña and the extraordinary year of 2011. *Journal of Geophysical Research: Biogeosciences*, *118*, 1247–1255. https://doi.org/10.1002/jgrg.20100

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823.

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

Bowman, K. W., Liu, J., Bloom, A. A., Parazoo, N. C., Lee, M., Jiang, Z., et al. (2017). Global and Brazilian carbon response to El Niño Modoki 2011–2010. *Earth and Space Science*, *4*, 637–660. https://doi.org/10.1002/2016EA000204

Brandt, M., Wigneron, J.-P., Chave, J., Tagesson, T., Penuelas, J., Ciais, P., et al. (2018). Satellite passive microwaves reveal recent climate-induced carbon losses in African drylands. *Nature Ecology & Evolution*, *2*(5), 827.

Buermann, W., Forkel, M., O'Sullivan, M., Sitch, S., Friedlingstein, P., Haverd, V., et al. (2018). Widespread seasonal compensation effects of spring warming on northern plant productivity. *Nature*, *562*(7725), 110.

Canadell, C. P., Sabine, C., & Joos, F. J. (2012). REgional Carbon Cycle Assessment and Processes (RECCAP). *Biogeosciences*, *9*, 2889–2904.

Chen, C.-T. A., Huang, T.-H., Chen, Y.-C., Bai, Y., He, X., & Kang, Y. (2013). Air-sea exchanges of CO_2 in the world's coastal seas. *Biogeosciences*, *10*(10), 6509–6544. http://www.biogeosciences.net/10/6509/2013/

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bron, F.-M., et al. (2005). Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data. *Journal of Geophysical Research*, *110*, D24309. http://doi.org/10.1029/2005jd006390

Chuvieco, E., Lizundia-Loiola, J., Pettinari, M. L., Ramo, R., Padilla, M., Tansey, K., et al. (2018). Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth System Science Data*, *10*(4), 2015–2031.

Chuvieco, E., Yue, C., Heil, A., Mouillot, F., Alonso-Canas, I., Padilla, M., et al. (2016). A new global burned area product for climate assessment of fire impacts. *Global Ecology and Biogeography*, *25*(5), 619–629.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., et al. (2014). Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system. *Biogeosciences*, *11*(13), 3547–3602. http://www.biogeosciences.net/11/3547/2014/

Clark, D., Mercado, L., Sitch, S., Jones, C., Gedney, N., Best, M., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description—Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, *4*(3), 701–722.

Di Gregorio, A. (2005). *Land cover classification system: Classification concepts and user manual: LCCS* (Vol. 8). Rome: Food & Agriculture Org.

Dlugokencky, E., & Tans, P. (2018). Trends in atmospheric carbon dioxide, National Oceanic & Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL). http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, last access: 4 September 2018.

ESA (2017). Land Cover CCI Product User Guide Version 2. Tech. Rep. (2017). Retrieved from http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf

FAO (2015). Global Forest Resources Assessment 2015 (FRA 2015). Rome: Food and Agriculture Organization of the United Nations.

FAOSTAT (2015). FAOSTAT: Food and Agriculture Organization of the United Nations. http://www.fao.org

Fan, Y., Miguez-Macho, G., Jobbágy, E. G., Jackson, R. B., & Otero-Casal, C. (2017). Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences*, *114*(40), 10,572–10,577.

Fernndez-Martnez, M., Sardans, J., Chevallier, F., Ciais, P., Obersteiner, M., Vicca, S., et al. (2019). Global trends in carbon sinks and their relationships with CO_2 and temperature. *Nature Climate Change*, *9*, 73–79. http://doi.org/10.1038/s41558-018-0367-7

Gaubert, B., Stephens, B. B., Basu, S., Chevallier, F., Deng, F., Kort, E. A., et al. (2019). Global atmospheric CO_2 inverse models converging on neutral tropical land exchange, but disagreeing on fossil fuel and atmospheric growth rate. *Biogeosciences*, *16*(1), 117–134. https://www.biogeosciences.net/16/117/2019/

Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, *178*, 31–41. http://doi.org/10.1016/j.rse.2016.02.054

Gloor, E., Wilson, C., Chipperfield, M. P., Chevallier, F., Buermann, W., Boesch, H., et al. (2018). Tropical land carbon cycle responses to 2015/16 El Niño as recorded by atmospheric greenhouse gas and remote sensing data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20170302.

Goll, D. S., Vuichard, N., Maignan, F., Jornet-Puig, A., Sardans, J., Violette, A., et al. (2017). A representation of the phosphorus cycle for ORCHIDEE (revision 934 3985). *Geoscientific Model Development Discussions*, *10*, 3745–3770.

Graven, H. D., Keeling, R. F., Piper, S. C., Patra, P. K., Stephens, B. B., Wofsy, S. C., et al. (2013). Enhanced seasonal exchange of CO_2 by northern ecosystems since 1960. *Science*, *341*(6150), 1085–1089. http://www.sciencemag.org/content/341/6150/1085.abstract

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853.

Hansis, E., Davis, S. J., & Pongratz, J. (2015). Relevance of methodological choices for accounting of land use change carbon fluxes. *Global Biogeochemical Cycles*, *29*, 1230–1246. https://doi.org/10.1002/2014GB004997

Harris, I., Jones, P. D., Osborn, T. J., & Lister, D. H. (2014). Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *International journal of climatology*, *34*(3), 623–642.

Hartmann, J., Jansen, N., Dürr, H. H., Kempe, S., & Köhler, P. (2009). Global CO_2-consumption by chemical weathering: What is the contribution of highly active weathering regions?. *Global and Planetary Change*, *69*(4), 185–194.

Haverd, V., Smith, B., Nieradzik, L., Briggs, P., Woodgate, W., Trudinger, C., et al. (2018). A new version of the cable land surface model (subversion revision r4601) incorporating land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to plant coordination of 956 photosynthesis. *Geoscientific Model Development*, *11*, 2995–3026. gmd-11-2995-2018.

Heinimann, A., Mertz, O., Frolking, S., Christensen, A. E., Hurni, K., Sedano, F., Chini, L. P., Sahajpal, R., Hansen, M., & Hurtt, H. (2017). A global view of shifting cultivation: Recent, current, and future extent. *PLOS One*, *12*(9), e0184479. https://doi.org/10.1371/journal.pone.0184479

Hooijer, A., Page, S., Canadell, J. G., Silvius, M., Kwadijk, J., Wsten, H., & Jauhiainen, J. (2010). Current and future CO_2 emissions from drained peatlands in Southeast Asia. *Biogeosciences*, *7*, 1505–1514. http://doi.org/10.5194/bg-7-1505-2010

Houghton, R. A., House, J. I., Pongratz, J., van der Werf, G. R., DeFries, R. S., Hansen, M. C., et al. (2012). Carbon emissions from land use and land-cover change. *Biogeosciences*, *9*(12), 5125–5142. http://www.biogeosciences.net/9/5125/2012/

Houghton, R., & Nassikas, A. A. (2017). Global and regional fluxes of carbon from land use and land cover change 1850–2015. *Global Biogeochemical Cycles*, *31*, 456–472. https://doi.org/10.1002/2016GB005546

Humphrey, V., Zscheischler, J., Ciais, P., Gudmundsson, L., Sitch, S., & Seneviratne, S. I. (2018). Sensitivity of atmospheric CO_2 growth rate to observed changes in terrestrial water storage. *Nature*, *560*(7720), 628.

Hurtt, G., Chini, L. P., Frolking, S., Betts, R., Feddema, J., Fischer, G., et al. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, *109*(1-2), 117–161.

Hurtt, G., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., et al. (2017). Harmonization of global land use scenarios (LUH2): Historical v2.1h. Earth System Grid Federation. https://doi.org/10.22033/ESGF/input4MIPs.1127

IPCC (2001). *Climate change 2001: The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change Edited by* J. T. Houghton, Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, & C. A. Johnson. Cambridge and New York: Cambridge University Press,.

IPCC (2007). *Climate change 2007: The Physical Science Basis (Vol. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Edited by* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, & H. L. Miller. Cambridge and New York: Cambridge University Press.

IPCC (2013). *Climate Change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Edited by* Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, B., & Midgley, B. M.Cambridge: Cambridge University Press.

Joetzjer, E., Delire, C., Douville, H., Ciais, P., Decharme, B., Carrer, D., et al. (2015). Improving the ISBA (CC) land surface model simulation of water and carbon fluxes and stocks over the Amazon forest. *Geoscientific Model Development*, *8*(6), 1709–1727.

Kato, E., Kinoshita, T., Ito, A., Kawamiya, M., & Yamagata, Y. (2013). Evaluation of spatially explicit emission scenario of land-use change and biomass burning using a process-based biogeochemical model. *Journal of Land Use Science*, *8*(1), 104–122.

Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S., Gerber, M., Doney, S. C., Graven, H. D., et al. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, *10*(4), 2169–2191. http://www.biogeosciences.net/10/2169/2013/

Klein Goldewijk, K., Beusen, A., van Drecht, G., & de Vos, M. (2011). The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography*, *20*(1), 73–86. https://doi.org/10.1111/j.1466-8238.2010.00587.x

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, *93*, 5–48. http://doi.org/10.2151/jmsj.2015-001

Krinner, G., Viovy, N., de Noblet-Ducoudríe, N., Ogée, J., Polcher, J., Friedlingstein, P., et al. (2005). A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochemical Cycles*, *19*, GB1015. http://doi.org/10.1029/2003GB002199

Landschützer, P., Gruber, N., Haumann, F. Alexander, Rdenbeck, C., Bakker, DorotheeC. E., van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., et al. (2018b). Global carbon budget 2018. *Earth System Science Data*, *10*(4), 2141–2194. http://doi.org/10.5194/essd-10-2141-2018

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., et al. (2018). Global carbon budget 2017. *Earth System Science Data*, *10*(1), 405–448.

Le Queré, C., Raupach, M. R., Canadell, J. G., & Marland et al., G. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836. https://doi.org/10.1038/ngeo689

Li, F., Bond-Lamberty, B., & Levis, S. (2014). Quantifying the role of fire in the Earth system—Part 2: Impact on the net carbon balance of global terrestrial ecosystems for the 20th century. *Biogeosciences*, *11*(5), 1345–1360.

Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., et al. (2018). Gross and net land cover changes based on plant functional types derived from the annual ESA CCI land cover maps. *Earth System Science Data*, *10*, 219–234.

Lienert, S., & Joos, F. (2018). A Bayesian ensemble data assimilation to constrain model parameters and land-use carbon emissions. *Biogeosciences*, *15*(9), 2909–2930.

Liu, X., Zhang, Y., Han, W., Tang, A., Shen, J., Cui, Z., et al. (2013). Enhanced nitrogen deposition over China. *Nature*, *494*(7438), 459.

Lucht, W., Prentice, I. C., Myneni, R. B., Sitch, S., Friedlingstein, P., Cramer, W., et al. (2002). Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science*, *296*(5573), 1687–1689.

Marland, T. A., Boden, G., & Andres, R. J. (2008). *Global, regional, and national fossil fuel CO_2 emissions. In Trends: A compendium of data on global change*. Oak Ridge, Tenn., U.S.A: Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy.

Mason, S. J., & Goddard, L. (2001). Probabilistic precipitation anomalies associated with ENSO. *Bulletin of the American Meteorological Society*, *82*(4), 619–638. https://doi.org/10.1175/1520-0477(2001)082<0619:PPAAWE>2.3.CO;2

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2018). Developments in the MPI-M Earth System Model version 1.2 (MPI1064 ESM 1.2) and its response to increasing CO_2. *Journal of Advances in Modeling Earth Systems*, *11*, 998–1038. https://doi.org/10.1029/2018MS001400

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H., Bouwman, A., et al. (2010). Global nutrient export from watersheds 2 (news 2): Model development and implementation. *Environmental Modelling & Software*, *25*(7), 837–853.

Meiyappan, P., Jain, A. K., & House, J. I. (2015). Increased influence of nitrogen limitation on CO_2 emissions from future land use and land use change. *Global Biogeochemical Cycles*, *29*, 1524–1548. https://doi.org/10.1002/2015GB005086

Melton, J., & Arora, V. (2016). Competition between plant functional types in the Canadian Terrestrial Ecosystem Model (ctem) v. 2.0. *Geoscientific Model Development*, *9*, 323–361.

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., & Cox, P. M. (2009). Impact of changes in diffuse radiation on the global land carbon sink. *Nature*, *458*(7241), 1014–1017. https://doi.org/10.1038/nature07949

Moore, S., Evans, C. D., Page, S. E., Garnett, M. H., Jones, T. G., Freeman, C., et al. (2013). Deep instability of deforested tropical peatlands revealed by fluvial organic carbon fluxes. *Nature*, *493*(7434), 660.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., et al. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, *300*(5625), 1560–1563. http://doi.org/10.1126/science.1082750

Oleson, K., Lawrence, D., Bonan, G., Drewniak, B., Huang, M., Koven, C., et al. (2013). Technical description of version 4.5 of the Community Land Model (CLM): NCAR Tech.

Olivier, J. G. J., Schure, K. M., & Peters, J. A. H. W. (2017). Trends in global CO_2 and total greenhouse gas emissions: PBL Netherlands Environmental Assessment Agency, 5.

Patra, P. K., Ishizawa, M., Maksyutov, S., Nakazawa, T., & Inoue, G. (2005). Role of biomass burning and climate anomalies for land-atmosphere carbon fluxes based on inverse modeling of atmospheric CO_2. *Global Biogeochemical Cycles*, *19*, GB3005. http://doi.org/10.1029/2004GB002258

Patra, P. K., Takigawa, M., Watanabe, S., Chandra, N., Ishijima, K., & Yamashita, Y. (2018). Improved chemical tracer simulation by MIROC4.0-based Atmospheric Chemistry-Transport Model (MIROC4-ACTM). *Sola*, *14*, 91–96.

Peters, G. P., Davis, S. J., & Andrew, R. (2012). A synthesis of carbon in international trade. *Biogeosciences*, *9*(8), 3247–3276. http://www.biogeosciences.net/9/3247/2012/

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., et al. (2013). Global atmospheric carbon budget: Results from an ensemble of atmospheric CO_2 inversions. *Biogeosciences*, *10*(10), 6699–6720. http://doi.org/10.5194/bg-10-6699-2013

Piao, S., Huang, M., Liu, Z., Wang, X., Ciais, P., Canadell, J. G., et al. (2018). Lower land-use emissions responsible for increased net land carbon sink during the slow warming period, *11*, 739–743. http://doi.org/10.1038/s41561-018-0204-7

Pongratz, J., Reick, C. H., Houghton, R., & House, J. (2014). Terminology as a key uncertainty in net land use and land cover change carbon flux estimates. *Earth System Dynamics*, *5*, 177–195.

Poulter, B., Cadule, P., Cheiney, A., Ciais, P., Hodson, E., Peylin, P., et al. (2015). Sensitivity of global terrestrial carbon cycle dynamics to variability in satellite-observed burned area: Fire and carbon-cycle feedbacks, *29*, 207–222. http://doi.org/10.1002/2013gb004655

Poulter, B., Frank, D., Ciais, P., Myneni, R. B., Andela, N., Bi, J., et al. (2014). Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature*, *509*, 600–603. https://doi.org/10.1038/nature13376

Poulter, B., Frank, D., Hodson, E., & Zimmermann, N. (2011). Impacts of land cover and climate data selection on understanding terrestrial carbon dynamics and the CO_2 airborne fraction. *Biogeosciences*, *8*, 2027–2026.

Pugh, ThomasA. M., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., & Calle, L. (2019). Role of forest regrowth in global carbon sink dynamics. *Proceedings of the National Academy of Sciences*, *116*, 4382–4387.

Quilcaille, Y., Gasser, T., Ciais, P., Lecocq, F., Janssens-Maenhout, G., & Mohr, S. (2018). Uncertainty in projected climate change arising from uncertain fossil-fuel emission factors. *Environmental Research Letters*, *13*, 44017. http://doi.org/10.1088/1748-9326/aab304

Randerson, J. T., Van Der Werf, G. R., Giglio, L., Collatz, G. J., & Kasibhatla, P. S. (2017). Global Fire Emissions Database, version 4.1 (GFEDv4). ORNL Distributed Active Archive Center. Retrieved from https://daac.ornl.gov/cgi-bin/dsviewer.pl?dsid=1293

Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., et al. (2013). Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geoscience*, *6*(8), 597.

Reick, C., Raddatz, T., Pongratz, J., & Claussen, M. (2010). Contribution of anthropogenic land cover change emissions to pre-industrial atmospheric CO_2. *Tellus B*, *62*, 329–336. http://doi.org/10.1111/j.1600-0889.2010.00479.x

Resplandy, L., Keeling, R., Rödenbeck, C., Stephens, B., Khatiwala, S., Rodgers, K., et al. (2018). Revision of global carbon fluxes based on a reassessment of oceanic and riverine carbon transport. *Nature Geoscience*, *11*(7), 504.

Rödenbeck, C, Houweling, S., Gloor, M., & Heimann, M. (2003). CO_2 flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics*, *3*(6), 1919–1964.

Rödenbeck, C., Zaehle, S., Keeling, R., & Heimann, M. (2018). History of El Niño impacts on the global carbon cycle 1957–2017: A quantification from atmospheric CO_2 data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20170303. https://doi.org/10.1098/rstb.2017.0303

Saeki, T., & Patra, P. K. (2017). Implications of overestimated anthropogenic CO_2 emissions on East Asian and global land CO_2 flux inversion. *Geoscience Letters*, *4*(1), 9.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlstrm, A., et al. (2015). Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, *12*(3), 653–679. http://www.biogeosciences.net/12/653/2015/

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., et al. (2013). Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades. *Biogeosciences Discussions*, *10*(12), 20,113–20,177. http://www.biogeosciences-discuss.net/10/20113/2013/

Smith, B., Warlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, *11*, 2027–2054.

Thomas, R. T., Prentice, I. C., Graven, H., Ciais, P., Fisher, J. B., Hayes, D. J., et al. (2016). Increased light-use efficiency in northern terrestrial ecosystems indicated by CO_2 and greening observations. *Geophysical Research Letters*, *43*, 11,339–11,349. https://doi.org/10.1002/2016GL070710

Tian, H., Chen, G., Lu, C., Xu, X., Hayes, D. J., Ren, W., et al. (2015). North American terrestrial CO_2 uptake largely offset by CH_4 and N_2O emissions: Toward a full accounting of the greenhouse gas budget. *Climatic Change*, *129*(3-4), 413–426.

Tian, H., Yang, J., Lu, C., Xu, R., Canadell, J., Jackson, R., et al. (2018). The global N_2O Model Intercomparison Project (NMIP). *Bulletin of the American Meteorological Society*, *99*, 1231–1251.

UNFCCC (2015). Adoption of the Paris Agreement. FCCC/CP/2015/L.9/Rev.1

UNFCCC (2018). National inventory submissions 2018, United Nations Framework Convention on Climate Change. Available at: https://unfccc.int/process/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/national-inventory-submissions-2018

Van Leeuwen, T., Peters, W., Krol, M., & van der Werf, G. (2013). Dynamic biomass burning emission factors and their impact on atmospheric CO mixing ratios. *Journal of Geophysical Research: Atmospheres*, *118*, 6797–6815. https://doi.org/10.1002/jgrd.50478

van Schaik, E., Killaars, L., Smith, N. E., Koren, G., van Beek, L., Peters, W., & van der Laan-Luijkx, I. T. (2018). Changes in surface hydrology, soil moisture and gross primary production in the Amazon during the 2015/2016 El Niño. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1760), 20180084.

van der Laan-Luijkx, I. T., van der Velde, I. R., van der Veen, E., Tsuruta, A., Stanislawska, K., Babenhauserheide, A., et al. (2017). The CarbonTracker Data Assimilation Shell (CTDAS) v1.0: Implementation and global carbon balance 2001–2015. *Geoscientific Model Development*, *10*, 2785–2800. http://doi.org/10.5194/gmd-10-2785-2017

Viovy, N. (2016). Cruncep data set. available at: ftp://nacp.ornl.gov/synthesis/2009/frescati/temp/land_use_change/original/readme.htm

Walker, A. P., Quaife, T., van Bodegom, P. M., De Kauwe, M. G., Keenan, T. F., Joiner, J., et al. (2017). The impact of alternative trait-scaling hypotheses for the maximum photosynthetic carboxylation rate (Vcmax) on global gross primary production. *New Phytologist*, *215*(4), 1370–1386.

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R. A., Nojiri, Y., et al. (2013). Global ocean carbon uptake: Magnitude, variability and trends. *Biogeosciences*, *10*, 1983–2000.

Wilkenskjeld, S., Kloster, S., Pongratz, J., Raddatz, T., & Reick, C. H. (2014). Comparing the influence of net and gross anthropogenic land-use and land-cover changes on the carbon cycle in the MPI-ESM. *Biogeosciences*, *11*, 4817–4828.

Yue, C., Ciais, P., Cadule, P., Thonicke, K., Archibald, S., Poulter, B., et al. (2014). Modelling the role of fires in the terrestrial carbon balance by incorporating spitfire into the global vegetation model ORCHIDEE—Part 1: Simulating historical global burned area and fire regimes. *Geoscientific Model Development*, *7*(6), 2747–2767.

Yue, C., Ciais, P., Cadule, P., Thonicke, K., & van Leeuwen, T. T. (2015). Modelling the role of fires in the terrestrial carbon balance by incorporating spitfire into the global vegetation model ORCHIDEE—Part 2: Carbon emissions and the role of fires in the global carbon balance. *Geoscientific Model Development*, *8*(5), 1321–1338. http://www.geosci-model-dev.net/8/1321/2015/

Zaehle, S., Friend, A. D., Friedlingstein, P., Dentener, F., Peylin, P., & Schulz, M. (2010). Carbon and nitrogen cycle dynamics in the O-CN land surface model: 2. Role of the nitrogen cycle in the historical terrestrial carbon balance. *Global Biogeochemical Cycles*, *24*, GB1006. http://doi.org/10.1029/2009GB003522

Zscheischler, J., Mahecha, M. D., Avitabile, V., Calle, L., Carvalhais, N., Ciais, P., et al. (2017). Reviews and syntheses: An empirical spatiotemporal description of the global surface-atmosphere carbon fluxes: Opportunities and data limitations. *Biogeosciences*, *14*(15), 3685–3703.