# ARTICLE

# Net carbon emissions from African biosphere dominate pan-tropical atmospheric $CO_2$ signal

Paul I. Palmer [1,2], Liang Feng[1,2], David Baker[3], Frédéric Chevallier [4], Hartmut Bösch[5,6] & Peter Somkuti[5,6]

Tropical ecosystems are large carbon stores that are vulnerable to climate change. The sparseness of ground-based measurements has precluded verification of these ecosystems being a net annual source (+ve) or sink (−ve) of atmospheric carbon. We show that two independent satellite data sets of atmospheric carbon dioxide ($CO_2$), interpreted using independent models, are consistent with the land tropics being a net annual carbon emission of $\left(\text{median}^{\text{maximum}}_{\text{minimum}}\right)$ $1.03^{+1.73}_{-0.20}$ and $1.60^{+2.11}_{+1.39}$ petagrams (PgC) in 2015 and 2016, respectively. These pan-tropical estimates reflect unexpectedly large net emissions from tropical Africa of $1.48^{+1.95}_{+0.80}$ PgC in 2015 and $1.65^{+2.42}_{+1.14}$ PgC in 2016. The largest carbon uptake is over the Congo basin, and the two loci of carbon emissions are over western Ethiopia and western tropical Africa, where there are large soil organic carbon stores and where there has been substantial land use change. These signals are present in the space-borne $CO_2$ record from 2009 onwards.

[1] School of GeoSciences, University of Edinburgh, Edinburgh EH9 3FF, UK. [2] National Centre for Earth Observation at the University of Edinburgh, Edinburgh EH9 3FF, UK. [3] Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins 80523-1375 Colorado, USA. [4] Laboratoire des Sciences du Climat et de l'Environnement/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91198 Gif-sur-Yvette, France. [5] Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK. [6] National Centre for Earth Observation at the University of Leicester, Leicester LE1 7RH, UK. Correspondence and requests for materials should be addressed to P.I.P. (email: paul.palmer@ed.ac.uk)

Tropical terrestrial ecosystems store large amounts of carbon in plants and soil, but are particularly vulnerable to changes in climate[1,2]. They release $CO_2$ via autotrophic and heterotrophic respiration and via fire, and take up $CO_2$ via photosynthesis. The terrestrial tropics, defined between 23.44°S and 23.44°N, include 30% of the global land surface and approximately a third of all Earth's three billion trees[3] and their stored carbon. Our knowledge of the tropical carbon budget has improved significantly over the past few decades mainly due to networks of sample plot measurements[4], micro-meteorological measurements of carbon fluxes of forest ecosystems[5], remote sensing of vegetation state or of land use change[6], and sparsely-distributed ground-based mole fraction measurements[7,8] of atmospheric $CO_2$. Despite these efforts, carbon fluxes from tropical ecosystems remain one of the largest uncertainties in the global carbon cycle[9,10] and impose a similar uncertainty on our ability to predict future climate change.

We use a range of global satellite data ("Methods") to study the carbon cycle over the tropics from 2009 to 2017 with a focus on 2015/2016 when two independent satellites were observing atmospheric $CO_2$. We use total column $CO_2$ dry air mole fraction ($X_{CO_2}$) retrievals from the Japanese Greenhouse gases Observing SATellite[11] (GOSAT) from mid-2009 until 2017 and from the NASA Orbiting Carbon Observatory[12] (OCO-2) from late 2014 to 2017. For comparative purposes, we use an inter-calibrated network of various mole fraction data[7,8] ("Methods"). We interpret these ground-based mole fraction and remotely-sensed column mole fraction data using three independent atmospheric transport models, driven by different a priori $CO_2$ flux estimates, and their counterpart inverse methods ("Methods"). The result is a range of geographically-resolved a posteriori $CO_2$ fluxes for the globe. We report our results over land as net biosphere fluxes ("Methods"), representing the net carbon flux exchange with the atmosphere from above-ground biomass and soils across sub-continental regions. To interpret these $CO_2$ fluxes in terms of the underlying land surface processes, we use correlative satellite data products (Methods): vegetation indices that provide information about leaf phenology[13]; changes in water storage[14]; a measure of photosynthesis[15]; and formaldehyde columns that provide information about the location and timing of fires[16]. We use dry matter (DM) burned estimates[17] inferred from remotely sensed land surface properties, and analysed meteorological fields of surface temperature and precipitation from the GEOS-5 (GEOS-FP) model ("Methods").

We use three inverse methods ("Methods") representing a range of atmospheric transport models, driving meteorology, and estimation methods. We focus on the GEOS-Chem atmospheric transport model[18] and discuss differences with the other models. Our primary study period is 2014–2017 when there is overlap between GOSAT and OCO-2 data, coinciding with the El Niño event[19–21]. The a posteriori global atmospheric growth rate of $CO_2$, inferred from ground-based data ("Methods") and converted from satellite-based flux estimates, ranges from 4.5 to 6.1 PgC year$^{-1}$ over our study, consistent with values inferred from the $CO_2$ mass inferred directly from the atmospheric mole fraction data multiplied by the total mass of dry air in the atmosphere.

Our analysis of the GOSAT and OCO-2 data reveals that the land tropics are a net annual $CO_2$ emission of (median$_{minimum}^{maximum}$) $1.03_{-0.20}^{+1.73}$ and $1.60_{+1.39}^{+2.11}$ petagrams (PgC) in 2015 and 2016, respectively, and larger than estimates inferred from changes in above-ground biomass[22–24]. The range of individual model estimates can be relatively large, particularly for regions where the net carbon budget is small, but nevertheless a coarse picture of the changing carbon budget emerges from our analysis. We find a robust signal over northern tropical Africa that is responsible for the majority of the pan-tropical net carbon signal, which cannot be explained by potential measurement or model biases. The largest seasonal uptake is over the northern Congo basin, as expected, and the largest emissions are found over western Ethiopia and western tropical Africa during March and April when it is hottest and driest. Although caution should be exercised when interpreting regions smaller than 1000 km, these emission focal points are a robust feature of the GOSAT record that starts in 2009. While we do not provide a definitive explanation for this seasonal signal, we argue that a comparatively small constant $CO_2$ flux, e.g., from soils due to sustained land degradation[25], could manifest as a seasonal net carbon source.

## Results

**Pan-tropical carbon flux estimates.** Figures 1a and d demonstrate that the sparse ground-based measurements provide insufficient information to determine robust estimates of tropical land carbon fluxes across the three groups, even on a pan-tropical scale. Differences in atmospheric model transport, assumptions about model errors, and differences between a priori land biosphere fluxes result in sometimes-inconsistent a posteriori estimates[9,10]. This has hampered the ability of the wider Earth system science community to understand large-scale responses of the carbon cycle to climate. On a broad scale, we can make two observations. First, we find that using column observations of $X_{CO_2}$ from GOSAT and OCO-2 results in more consistent a posteriori $CO_2$ flux estimates over the tropics (Fig. 1b–f), with a smaller inter-model spread of estimates[26], and a better agreement on the phase of the seasonal cycle than using only in situ observations of $CO_2$ (Fig. 1a–d). Second, the amplitude of the seasonal cycle of a posteriori $CO_2$ fluxes over the northern and southern tropical lands inferred by the satellite data is generally much larger than that inferred from the in situ data (Fig. 1b–e), with the exception of LSCE that is driven by a priori fluxes from the ORCHIDEE model ("Methods"). Differences between the amplitude of the seasonal cycle inferred by GEOS-Chem using GOSAT and OCO-2 data (Fig. 1b–e) are smaller than those from different models that use the same data (Fig. 1c–f). Assumptions about data analysis therefore still play a role in the a posteriori flux estimates, but these inter-model differences are generally small compared to differences between the a posteriori and a priori flux estimates. Together, these two observations suggest that the satellite data contain substantial information about the carbon cycle. For completeness, we refer the reader to Supplementary Figs. 1–6 and Supplementary note 1 for analysis and discussion of a posteriori $CO_2$ fluxes from all other regions across the world.

Table 1 shows that the land tropics are a net annual carbon emission of $1.03_{-0.20}^{+1.73}$ and $1.60_{+1.39}^{+2.11}$ PgC in 2015 and 2016, respectively, and larger than estimates inferred from changes in above-ground biomass[22–24]. We find that northern tropical fluxes are $1.54_{-0.12}^{+1.58}$PgC in 2015 that increase in 2016 to $1.72_{+1.61}^{+2.42}$PgC. Southern tropical fluxes are $-0.26_{-0.55}^{+0.23}$PgC in 2015 and $-0.18_{-0.41}^{-0.01}$PgC in 2016. Even on a pan-tropical scale for 2015 and 2016, reaching a consensus on the sign of the land flux (except for GOSAT ACOS data, Table 1) and on its seasonal amplitude ($1.39_{-0.20}^{+2.11}$PgC) represents a significant step forward for the carbon cycle community. For our analysis we have not quantified anomalous fluxes during El Niño[20,21].

**Continental-scale tropical carbon flux estimates.** During 2015 we find that net fluxes from tropical South America are $-0.26_{-0.58}^{+0.04}$PgC, tropical African fluxes are $1.48_{+0.80}^{+1.95}$PgC, and from
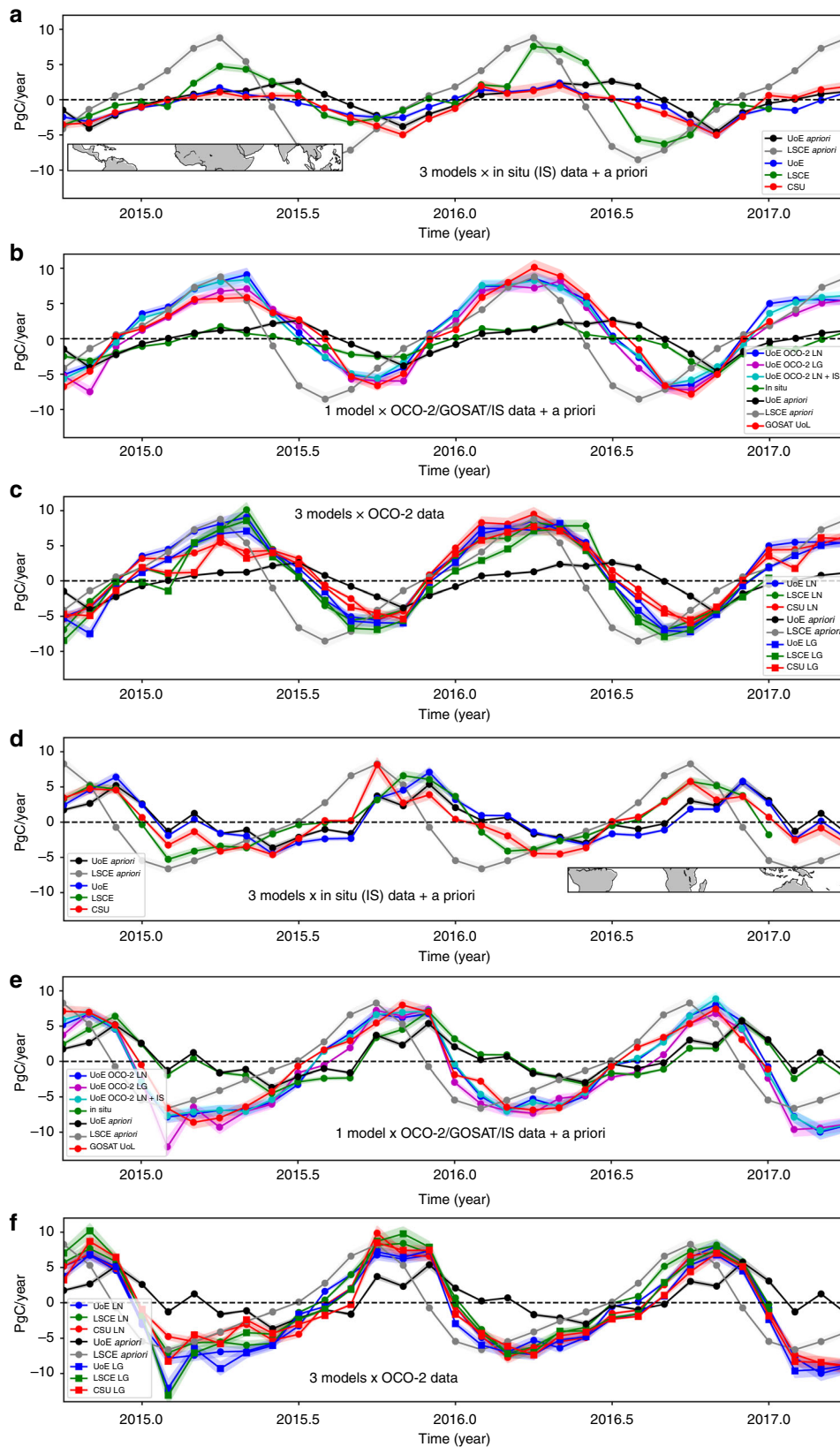
**Fig. 1** Northern and southern tropical $CO_2$ fluxes. Monthly a priori and a posteriori $CO_2$ fluxes (expressed as PgC year$^{-1}$, mid-2014 to 2017) from the (**a–c**) Northern and (**d–f**) Southern tropics inferred from (**a, d**) in situ mole fraction measurements and from (**b–c, e–f**) GOSAT and OCO-2 satellite measurements of $X_{CO2}$. Positive fluxes are from the land surface to the atmosphere. LN and LG denote $X_{CO2}$ measurements taken using nadir and glint observing modes, respectively. The geographical regions are shown inset of each upper panel

**Table 1 A priori and a posteriori net biosphere CO$_2$ flux estimates (PgC year$^{-1}$, positive values are to the atmosphere) for 2015 and 2015 for individual geographical regions over the tropics inferred from GOSAT and OCO-2 X$_{CO_2}$ retrievals**

| | 2015 | | | | | | 2016 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A priori | GOSAT | | OCO-2 | | | A priori | GOSAT | | OCO-2 | | |
| | | UoE/ACOS | UoE/UoL | UoE | LSCE | CSU | | UoE/ACOS | UoE/UoL | UoE | LSCE | CSU |
| Northern Tropical South America | −0.08 | −0.27 | 0.17 | 0.22 | 0.23 | −0.05 | 0.00 | −− | 0.17 | 0.16 | 0.39 | −0.06 |
| Southern Tropical South America | −0.01 | −0.31 | −0.50 | −0.31 | −0.19 | −0.21 | 0.35 | −− | −0.06 | 0.12 | 0.14 | −0.15 |
| Tropical South America | −0.09 | −0.58 | −0.33 | −0.09 | 0.04 | −0.26 | 0.35 | −− | 0.11 | 0.28 | 0.53 | −0.21 |
| Northern Tropical Africa | −0.05 | 0.60 | 0.69 | 1.89 | 1.54 | 1.91 | 0.06 | −− | 1.45 | 2.02 | 1.34 | 2.54 |
| Southern Tropical Africa | −0.13 | 0.31 | 0.11 | −0.36 | −0.06 | 0.04 | −0.16 | −− | 0.05 | −0.23 | −0.20 | −0.12 |
| Tropical Africa | −0.18 | 0.91 | 0.80 | 1.53 | 1.48 | 1.95 | −0.10 | −− | 1.50 | 1.79 | 1.14 | 2.42 |
| Northern Tropical Asia | 0.04 | −0.45 | 0.19 | −0.53 | −0.22 | −0.32 | 0.16 | −− | 0.18 | −0.54 | −0.12 | −0.06 |
| Southern Tropical Asia | 0.18 | 0.00 | 0.21 | 0.26 | 0.09 | 0.42 | 0.03 | −− | 0.11 | 0.14 | 0.10 | 0.10 |
| Tropical Asia | 0.22 | −0.45 | 0.40 | −0.27 | −0.13 | 0.10 | 0.19 | −− | 0.29 | −0.40 | -0.02 | 0.04 |
| Tropical Australia | −0.02 | −0.08 | −0.33 | −0.14 | −0.10 | −0.06 | 0.01 | −− | −0.51 | −0.08 | -0.05 | −0.14 |
| | | | | | | | | | | | | |
| Tropics | −0.07 | −0.20 | 0.54 | 1.03 | 1.29 | 1.73 | 0.45 | −− | 1.39 | 1.59 | 1.60 | 2.11 |
| Northern Tropics | −0.09 | −0.12 | 1.05 | 1.58 | 1.55 | 1.54 | 0.22 | −− | 1.80 | 1.64 | 1.61 | 2.42 |
| Southern Tropics | 0.02 | −0.08 | −0.51 | −0.55 | −0.26 | 0.19 | 0.23 | −− | −0.41 | −0.05 | −0.01 | −0.31 |

− denotes absence of data. UoL and ACOS denote independent retrieval data products from the University of Leicester and the NASA Atmospheric CO$_2$ Observation from Space, respectively. UoE, LSCE, and CSU denote the modelling groups from the University of Edinburgh, Laboratoire des Sciences du Climat et de l'Environnement, and Colorado State University, respectively. Colours are used to distinguish column/row groups

tropical Asia and tropical Australia are $-0.13^{+0.40}_{-0.45}$PgC and $-0.10^{-0.06}_{-0.33}$PgC, respectively (Table 1). In comparison, during 2016 tropical South America fluxes are $0.20^{+0.53}_{-0.21}$PgC, tropical African fluxes are $1.65^{+2.42}_{+1.14}$PgC, and tropical Asia and tropical Australia are $-0.01^{+0.29}_{-0.40}$PgC and $-0.11^{-0.05}_{-0.51}$PgC, respectively. The range of individual model estimates can be relatively large, particularly for regions where the net carbon budget is small, but nevertheless a coarse picture of the changing carbon budget emerges from our analysis (Table 1).

**Carbon flux estimates for northern tropical Africa and southern tropical South America.** Figures 2 and 3 shows carbon budgets for two contrasting tropical regions: southern tropical South America and northern tropical Africa. To explore the ability of these satellite data to constrain fluxes on smaller spatial scales, we present our results also as latitude-mean Hovmöller plots, reflecting that physical climate variations over the tropics are typically oriented E-W. In the absence of independent CO$_2$ data to evaluate these distributions, we interpret the a posteriori CO$_2$ fluxes using correlative satellite observations (Fig. 4).

Over southern tropical South America (Fig. 2a), UoE a posteriori fluxes are shifted from the a priori seasonal cycle, resulting in a better agreement with fluxes inferred from the same data using different models (Fig. 2a, bottom panel). A posteriori flux estimates inferred from GOSAT lie between a priori values and the fluxes inferred from OCO-2, reflecting the superior data density of OCO-2; fluxes inferred from GOSAT are insignificantly different from a priori values during early 2016 due to a very low density of measurements during this period. Differences in the spatial and temporal CO$_2$ flux distributions (Fig. 3a) demonstrate current limitations in our ability to infer spatial distributions of CO$_2$ fluxes[26,27]. We find that the a posteriori distributions of carbon flux over the El Niño period resemble the E-W dipole pattern of water storage (Fig. 4a), with larger positive (negative) anomalies towards the east (west) corresponding to larger positive (negative) CO$_2$ fluxes. The El Niño period also saw anomalous fire activity in the 2015 dry season (Fig. 4a) that reflects anomalous high temperatures and drought conditions, which increase the susceptibility of vegetation to ignite.

We find that GOSAT and OCO-2 X$_{CO_2}$[28] data consistently assign the largest seasonal cycle of carbon fluxes over the tropics to northern tropical Africa (Fig. 2b and 3b) with that region being responsible for the unexpectedly large pan-tropical net source of carbon (Table 1, Fig. 1). Over this region, we find close agreement between the a posteriori flux estimates on small spatial and temporal scales (Fig. 3b). The largest seasonal uptake is over the northern Congo basin, as expected, and the largest emissions are found over western Ethiopia and western tropical Africa during March and April when it is hottest and driest (Supplementary Figs. 7–11; Supplementary note 2). Although caution should be exercised when interpreting regions smaller than 1000 km, these emission focal points are a robust feature of our analysis that extends back through the GOSAT record to 2009. We do not rule out a role for regional systematic retrieval errors[29], but comparison to sparse independent data (Supplementary Fig. 12; Supplementary note 2) and the results from extensive sensitivity experiments (Supplementary Figs. 13–16; Supplementary note 3) support our results. The magnitude and approximate timing of the inferred seasonal cycle of net fluxes is consistent with the ORCHIDEE land surface model (Fig. 2b; Supplementary Discussion), although the model has larger uptake later in the year.

**Discussion**
Compared to tropical South America there is a lower baseline for precipitation, water storage, leaf phenology, and SIF over tropical North Africa (Fig. 4b), but there is a large seasonal cycle in temperature. We find a comparatively muted seasonal cycle of HCHO columns, but a much larger seasonal cycle of DM burned (Fig. 4b), which is due to predominant grassland fuel not producing sufficient energy to be directly lofted above the boundary layer where it can be observed as HCHO. For completeness, Supplementary Figs. 17–22 show similar Hovmöller plots but for all studied land regions for the GOSAT record from 2009 to 2017. Supplementary Fig. 23 shows regionally-mean values of SIF from 2009 to 2017.
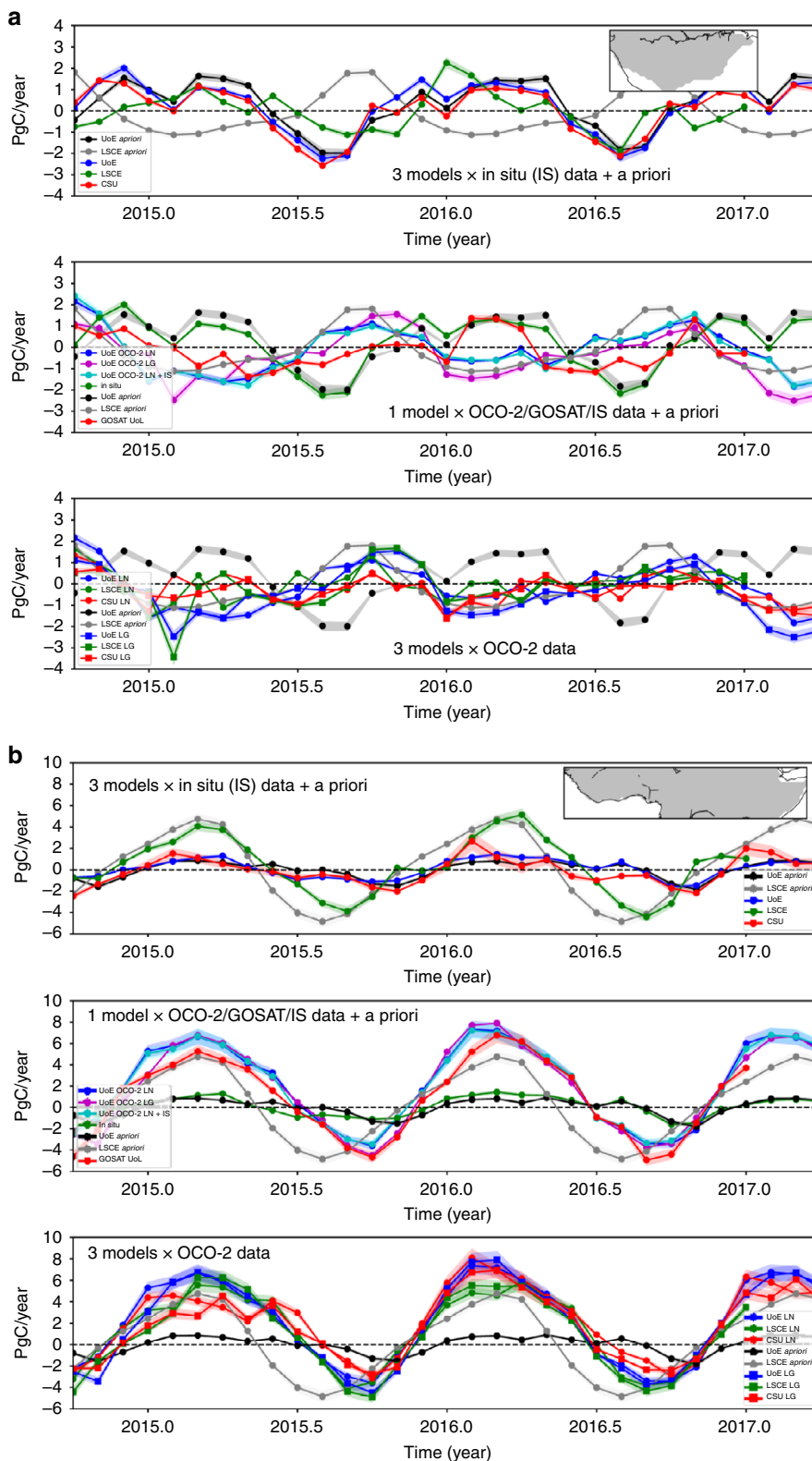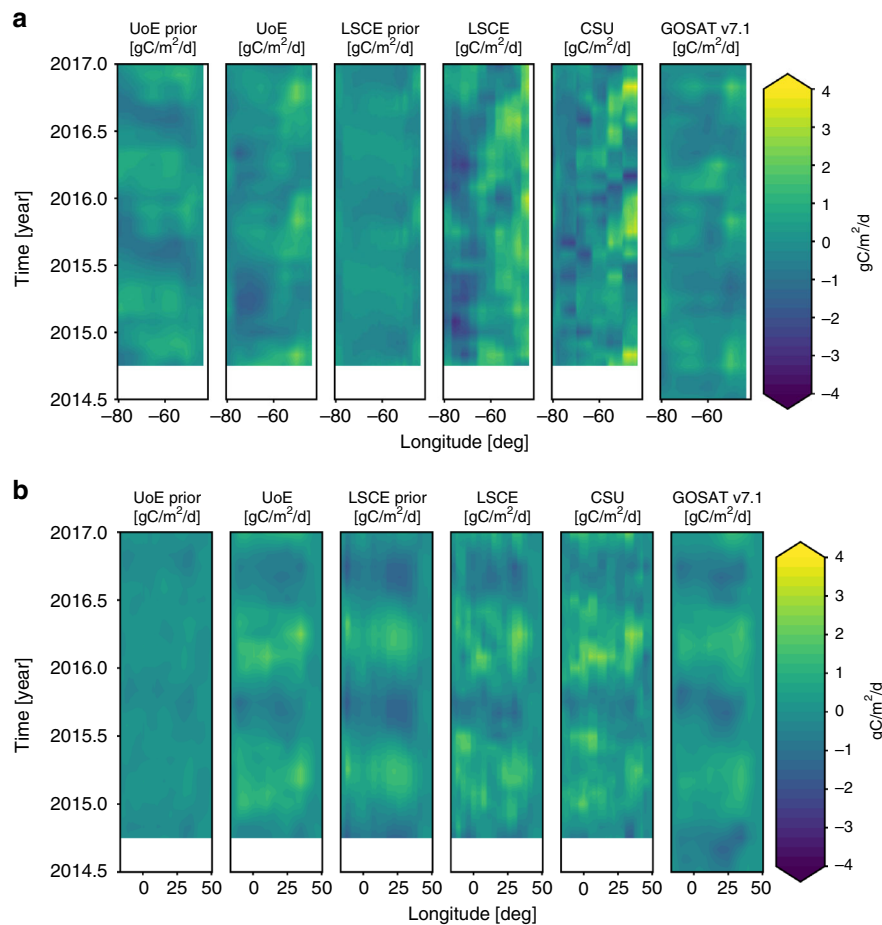
**Fig. 2** $CO_2$ fluxes from tropical South America and tropical northern Africa. Monthly a priori and a posteriori $CO_2$ fluxes (expressed as PgC year$^{-1}$, mid-2014 to 2017) from **a** southern tropical South America and **b** northern tropical Africa, inferred from in situ mole fraction measurements and from GOSAT and OCO-2 satellite measurements of $X_{CO2}$. Positive fluxes are from the land surface to the atmosphere. LN and LG denote $X_{CO2}$ measurements taken from nadir and glint observing modes, respectively. The geographical regions are shown inset of each upper panel.

**Fig. 3** $CO_2$ fluxes from tropical South America and tropical northern Africa. Hovmöller plots of monthly a priori and a posteriori $CO_2$ fluxes (expressed as gC m$^{-2}$ d$^{-1}$, mid 2014 to 2017), averaged over the latitude domain, from **a** southern tropical South America and **b** northern tropical Africa, inferred from GOSAT and OCO-2 satellite measurements of $X_{CO_2}$ taken from the nadir observing mode. Positive fluxes are from the land surface to the atmosphere.

Water storage records that start in 2002 reveal successive years of drought over this region (from 2009 in Supplementary Fig. 19) that could have impacted photosynthesis[15], land-use change[22,23], burning extent, and possibly soil carbon stocks[24]. Fire cannot explain these emissions (Supplementary Discussion), although it has a consistent seasonal cycle (Fig. 4b). Seasonally low soil water content will limit the source from soil microbial respiration, but even a small diffuse $CO_2$ flux from soils due to sustained land degradation[25] could manifest as a seasonal net carbon source (Supplementary Discussion).

We anticipate that our findings will help re-prioritise decadal science challenges for the carbon cycle community, particularly in the context of the Paris Agreement that implicitly relies on the continued operation of natural carbon sinks. Ultimately, deeper insights into the tropical carbon cycle will only be achieved by improved integration of in situ and remote-sensed data, for the short timescales, and pan-tropical sample plot data for the longer timescales.

## Methods

**In situ $CO_2$ mole fraction observations**. We use discrete (weekly) air samples from 105 sites and continuous (hourly) observations from 52 sites that are part of the global atmospheric surface $CO_2$ observations network. These were taken from the Observation Package (ObsPack) obspack_co2_1_GLOBALVIEW-plus_v2.1_2016_09_02 data product[7] for 2015, and from obspack_co2_1_NRT_v3.3_2017–04–19 for 2016–2017[8]; both datasets are produced by the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL).

**Satellite observations of column $CO_2$**. We use $X_{CO_2}$ data retrieved from the Japanese Greenhouse gases Observing SATellite (GOSAT) and the NASA Orbiting Carbon Observatory-2 (OCO-2). GOSAT[11] was launched in January 2009 in a sun-synchronous orbit with an equatorial crossing time of 1300. We use two independent GOSAT $XCO_2$ data products: v7.1 full-physics retrievals from the University of Leicester[30] (UoL), and B7.3 of the NASA Atmospheric $CO_2$ Observations from Space (ACOS[31]) activity. We use 10-s averages of the bias-corrected $X_{CO_2}$ B7.1r data product[32] over land from OCO-2 that is the current version used by the OCO-2 science team.[33,34]

**Enhanced Vegetation Index**. The Enhanced Vegetation Index (EVI) is a composite property of leaf area, chlorophyll and canopy structure[35]. We use MOD13C2 (MODIS/Terra Vegetation Indices Monthly L3 Global 0.05° CMG V006)[36] to get EVI information. The data are only retained with pixel reliability values masked as good data (0) or marginal data (1).

**Gravity recovery and climate experiment**. The Gravity Recovery and Climate Experiment (GRACE) provides information about changes in the water column[37–39]. Rooting depths of tropical terrestrial ecosystems will likely be sufficiently deep that we cannot establish a direct and immediate relationship between vegetation and changes in precipitation. Changes in gravity, due to changes in water column depth, provide a much stronger relationship with vegetation access to water. We use the surface mass change data based on the RL05 spherical harmonics from CSR (Center for Space Research at University of Texas, Austin), JPL (Jet Propulsion Laboratory) and GFZ (GeoforschungsZentrum Potsdam). The three different processing groups chose different parameters and solution strategies when deriving month-to-month gravity field variations from GRACE observations. We use the ensemble mean of the three data fields and multiply the data by the provided scaling grid. Data are available from http://grace.jpl.nasa.gov.

**Formaldehyde columns**. Formaldehyde (HCHO) columns are from the Ozone Monitoring Instrument[40] (OMI) aboard the NASA Aura satellite, which was
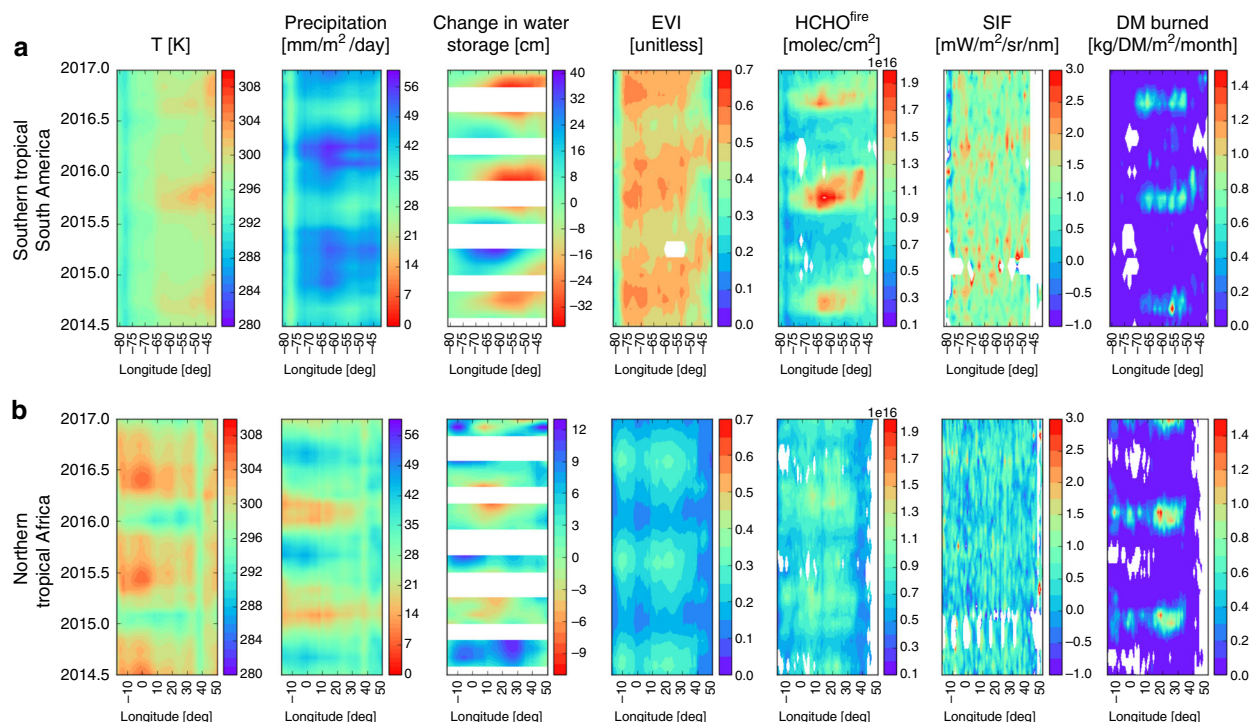
**Fig. 4** Correlative data to interpret regional $X_{CO_2}$ data. Correlative data over **a** southern tropical South America and **b** northern tropical Africa. The panels are from left to right: surface temperature (K), precipitation (mm m$^{-2}$ day$^{-1}$), water storage (cm), elevated vegetation index (m$^2$ m$^{-2}$), HCHO columns (molec cm$^{-2}$) filtered for fire activity using MODIS fire counts, and dry matter (DM) burned (kg DM$^{-1}$ m$^{-2}$ month$^{-1}$)

launched in a sun-synchronous orbit in 2009. We use the NASA OMHCHOv003 data product[16] from the NASA Data and Information Services Center, which fits HCHO slant columns in the 328.5–356.5 nm window and accounts for competing absorbers, the Ring effect, and undersampling. HCHO is a high-yield product of hydrocarbon oxidation[41,42]. It is also emitted as a direct emission from incomplete combustion[43,44]. We use the active fire data product[45] from the NASA Moderate Resolution Imaging Spectrometer (MODIS), derived from surface thermal IR anomalies, to isolate the pyrogenic HCHO signal.

**Satellite observations of solar induced fluorescence.** Satellite observations of solar induced fluorescence (SIF) are retrieved by the UoL from the GOSAT instrument[46]. SIF is a by-product of plant pigments absorbing incoming sunlight as part of photosynthesis. Of the solar radiation absorbed, ~20% is eventually dissipated as heat and typically <1–2% is emitted by SIF in the range 650–800 nm, peaking at 685–690 nm and 730–740 nm. GOSAT fits estimates of SIF at 755 nm[47]. We use the GOSAT SIF data product as a crude measure of photosynthetic capacity of regional ecosystems. We use a physically based retrieval scheme[47] with a focus on the bias correction procedure. We use a two-stage method. First, we isolate GOSAT measurements over non-vegetated areas using the ESA CCI Land Cover product V2.0.7[48] at 300 m resolution. Second, we apply a bias correction as an explicit function of time to ensure that instrumental effects are accounted for the entire date range of the SIF product.

**DM burned estimates.** DM burned estimates are taken from the Global Fire Emission Database[49] (GFED4). These estimates were derived by combined by satellite remote sensing observation of burned area and active fire data from MODIS.

**Atmospheric transport models and inverse methods.** To describe the relationship between surface fluxes of $CO_2$ and atmospheric $CO_2$ we use three atmospheric transport models: (1) GEOS-Chem global 3-D chemistry transport model[50,51] v9.02; (2) GSFC parameterised chemistry and transport model[52] (PCTM), and (3) Laboratoire de Météorologie Dynamique (LMDZ), version LMDZ3[53].

We run GEOS-Chem with a horizontal resolution of 4° (latitude) × 5° (longitude), driven by the GEOS-5 meteorological analyses (GEOS-FP from 2013) from the Global Modeling and Assimilation Office (GMAO) Global Circulation Model based at NASA Goddard Space Flight Center. We run the model using 47 vertical terrain-following sigma-levels that describe the atmosphere from the surface to 0.01 hPa, of which about 30 are typically below the dynamic tropopause.

We use well-established emission inventories as our a priori flux estimates: (1) weekly biomass burning emissions[49]; (2) monthly fossil fuel emissions[54,55]; (3) monthly climatological ocean fluxes[56]; and (4) three-hourly terrestrial biosphere fluxes[57].

The GEOS-Chem model uses an ensemble Kalman Filter (EnKF) framework[18,58] to infer $CO_2$ fluxes from the ground-based or space-based measurements of atmospheric $CO_2$. We use a total of 792 basis functions per month, split between 317 oceanic regions and 475 land regions. These regions are subdivisions of the 22 regions used in TransCom-3[9]. We assume a 50% uncertainty for monthly land terrestrial fluxes, and 40% for monthly ocean fluxes[49]. We assume land (ocean) a priori fluxes are correlated with a correlation length of 500 (800) km. We assume no observation error correlations, but include an additional 1.5 ppm uncertainty to the reported observation errors to account for model transport errors. We determine the terrestrial biosphere flux by subtracting the fossil fuel and cement production emission estimate (FF). This is a common approach[10,18,59], based on the assumption knowledge of FF flux is much better than that of the natural fluxes from the land and ocean.

The LMDZ model is run using a regular horizontal resolution of 3.75° (longitude) and 1.875° (latitude), with 39 hybrid layers in the vertical. Winds are nudged towards the 6-hourly ECMWF reanalysis[60] with a relaxation time of three hours. Fossil fuel burning emissions from the ODIAC model[54,55], including diurnal and day-of-week variability[61]. We also use monthly ocean fluxes[56], three-hourly biomass burning emissions (GFED 4.1 s until 2015 and GFAS afterwards), and climatological three-hourly biosphere-atmosphere fluxes taken as the 1989–2010 of a simulation of the ORganizing Carbon and Hydrology In Dynamic EcosystEms model (ORCHIDEE)[62], version 1.9.5.2.

The LMDZ CAMS inversion tool currently generates the global $CO_2$ atmospheric inversion product of the Copernicus Atmosphere Monitoring Service[63,64]. The minimum of the Bayesian cost function of the inversion problem is found by an iterative process using the Lanczos version of the conjugate gradient algorithm[65]. The inferred fluxes are estimated at each horizontal grid point of the transport model with a temporal resolution of eight days, separately for day-time and night-time. The state vector of the inversion system is therefore made of a succession of global maps with 9200 grid points. Per month it gathers 73,700 variables (four day-time maps and four night-time maps). It also includes a map of the total $CO_2$ columns at the initial time step of the inversion window in order to account for the uncertainty in the initial state of $CO_2$. Over land, the errors of the prior biosphere-atmosphere fluxes are assumed to dominate the error budget and the covariances are constrained by an analysis of mismatches with in situ flux measurements: temporal correlations on daily mean net carbon exchange (NEE) errors decay exponentially with a length of one month but night-time errors are assumed to be uncorrelated with daytime errors; spatial correlations decay

exponentially with a length of 500 km; standard deviations are set to 0.8 times the climatological daily-varying heterotrophic respiration flux simulated by ORCHIDEE with a ceiling of 4 gC/m$^2$/day. Over a full year, the total 1-sigma uncertainty for the prior land fluxes amounts to about 3.0 GtC/yr. The error statistics for the open ocean correspond to a global air-sea flux uncertainty about 0.5 GtC/yr and are defined as follows: temporal correlations decay exponentially with a length of one month; unlike land, daytime and night-time flux errors are fully correlated; spatial correlations follow an e-folding length of 1000 km; standard deviations are set to 0.1 gC/m$^2$/day. Land and ocean flux errors are not correlated.

PCTM is run at a horizontal resolution of 2.0° (latitude) × 2.5° (longitude) with 40 hybrid sigma levels in the vertical, driven by winds, surface pressure, and vertical mixing parameters from NASA MERRA2 reanalyses[66]. A priori fluxes for gross primary productivity, gross respiration, wildfires and biofuel emissions are taken from CASA-GFED3 land biosphere model[49,67,68]. Fossil fuel burning emissions from the ODIAC model[54,55], including diurnal and day-of-week variability[61], and air-sea $CO_2$ fluxes from three different sources: the NASA Ocean and Biosphere Model (NOBM)[69], and two $CO_2$ climatological flux products[56,70].

The CSU inversion scheme uses a variational data assimilation approach[71,72]. A priori $CO_2$ fluxes are run forward through PCTM at a 2.0° × 2.5° (lat/lon) resolution, with the resulting model-measurement residuals used in a 6.7° × 6.7° version of PCTM to estimate weekly flux corrections (no day/night split); no correlations in space or time are assumed. This configuration results in 54 × 27 × 4.33 ≈ 6300 monthly flux corrections being solved. The adjoint of PCTM, forced with the measurement mismatches, generates the gradient to the Bayesian cost function; this is used in a BFGS approach (pre-conditioned with the a priori flux uncertainties) to descend to the minimum, giving the optimal fluxes.

## Data availability
GOSAT V7.1 and SIF data are available from University of Leicester. OCO-2 retrievals were produced by the OCO-2 project at the Jet Propulsion Laboratory, California Institute of Technology, and obtained from the OCO-2 data archive maintained at the NASA Goddard Earth Science Data and Information Services Center. All correlative data are also freely available from NASA data repositories. The NOAA in situ data are freely available from the ESRL website (https://www.esrl.noaa.gov/gmd/ccgg/trends/data.html).

## Code availability
The community-led GEOS-Chem model of atmospheric chemistry and model is maintained centrally by Harvard University (http://acmg.seas.harvard.edu/geos/), and is available on request. The ensemble Kalman filter code is publicly available as PyOSSE (https://www.nceo.ac.uk/data-tools/atmospheric-tools/). The CAMS inversion system is available on simple request from FC. For access to PCTM, please contact DB.

## References
1. Lewis, S. L. et al. The 2010 Amazon Drought. *Science,* **331**, 554 (2011).
2. Page, S. E. et al. The amount of carbon released from peat and forest fires in Indonesia during 1997. *Nature* **420**, 61–65 (2002).
3. Crowther et al. Mapping tree density at a global scale. *Nature* **525**, 201–205 (2015).
4. Mahli et al. An international network to monitor the structure, composition, and dynamics of Amazonian forests (RAINFOR). *J. Veg. Sci.* **13**, 439–450 (2002).
5. Kruijt et al. The robustness of eddy correlation fluxes for Amazon rain forest conditions. *Ecol. Appl.* **14**, 101–113 (2004).
6. Yu, Y. & Saatchi, S. Sensitivity of L-band SAR backscatter to aboveground biomass of global forests. *Remote Sens.* **8**, 522–540 (2016).
7. Cooperative Global Atmospheric Data Integration Project, Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2015; obspack_co2_1_GLOBALVIEWplus_v2.1_2016_09_02; NOAA Earth System Research Laboratory, Global Monitoring Division. https://doi.org/10.15138/G3059Z (2016).
8. Carbontracker Team, Compilation of near real time atmospheric carbon dioxide data provided by NOAA and EC; obspack_co2_1_NRT_v3.3_2017-04-19; NOAA Earth System Research Laboratory, Global Monitoring Division. https://doi.org/10.15138/G3G01J (2017).
9. Gurney, K. et al. Towards robust regional estimates of $CO_2$ sources and sinks using atmospheric transport models. *Nature* **415**, 626–630 (2002).
10. Peylin, P. et al. Global atmospheric carbon budget: results from an ensemble of atmospheric $CO_2$ inversions. *Biogeosciences* **10**, 6699–6720 (2013).
11. Kuze, A., Suto, H., Nakajima, M. & Hamazaki, T. Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring. *Appl. Opt.* **35**, 6716–6733 (2009).
12. Eldering, A. et al. The Orbiting Carbon Observatory-2: first 18 months of science data products. *Atmos. Meas. Tech.* **10**, 549–563 (2017).
13. Huete, A. et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **83**, 195–213 (2002).
14. Tapley, B. D., Bettadpur, S., Watkins, M. & Reigber, C. The gravity recovery and climate experiment: mission overview and early results. *Geophys. Res. Lett.* **31**, L09607 (2004).
15. Frankenberg, C. et al, Remote sensing of solar induced chlorophyll fluorescence from satellites, airplanes and ground-based stations, *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).* https://doi.org/10.1109/IGARSS.2016.7729436 (2016).
16. González Abad, G. et al. Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval. *Atmos. Meas. Tech.* **8**, 19–32 (2015).
17. Randerson, J. T. et al. Global burned area and biomass burning emissions from small fires. *J. Geophys. Res.* **117**, G04012 (2012).
18. Feng, L. et al. Consistent regional fluxes of $CH_4$ and $CO_2$ inferred from GOSAT proxy XCH$_4$: XCO$_2$ retrievals, 2010–2014. *Atmos. Chem. Phys.* **17**, 4781–4797 (2017).
19. L'Heureux, M. L. et al, Observing and Predicting the 2015/16 El Niño. *Bull. Amer. Meteor. Soc.* https://doi.org/10.1175/BAMS-D-16-0009.1 (2017).
20. Liu, J. et al. Contrasting carbon cycle responses of the tropical continents to the 2015–2016 El Niño. *Science* **358**, eaam5690 (2017).
21. Patra, P. K. et al. The Orbiting Carbon Observatory (OCO-2) tracks 2–3 peta-gram increase in carbon release to the atmosphere during the 2014–2016 El Niño. *Sci. Rep.* **7**, 13567 (2017).
22. Baccini et al. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science* **358**, 230–234 (2017).
23. Brandt, M. et al. Satellite passive microwaves reveal recent climate-induced carbon losses in African drylands. *Nat. Ecol. Evol.* **2**, 827–835 (2018).
24. Borrelli, P. et al. An assessment of the global impact of 21$^{st}$ century land use change on soil erosion. *Nat. Commun.* **8**, 2013 (2017).
25. The IPBES assessment report on land degradation and restoration. Montanarella, L., Scholes, R., and Brainich, A. (eds). Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany (2018).
26. Basu, S. et al. The impact of transport model differences on $CO_2$ surface flux estimates from OCO-2 retrievals of column average $CO_2$. *Atmos. Chem. Phys.* **18**, 7189–7215 (2018).
27. Miller et al. Characterizing biosphere carbon balance using $CO_2$ observations from the OCO-2 satellite. *Atmos. Chem. Phys.* **18**, 6785–6799 (2018).
28. Hakkarainen et al. Direct space-based observations of anthropogenic $CO_2$ emission areas from OCO-2. *Geophys. Res. Lett.* **43**, 11400–11406 (2016).
29. O'Dell, C. W. et al. Improved Retrievals of Carbon Dioxide from the Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm. *Atmos. Meas. Tech.* **11**, 6539–6576 (2018).
30. Cogan, A. J. et al. Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): comparison with ground-based TCCON observations and GEOS-Chem model calculations. *J. Geophys. Res. Atmos.* **117**, D21 (2012).
31. OCO-2 Science Team/Michael Gunson, Annmarie Eldering (2016), ACOS GOSAT/TANSO-FTS ACOS Level 2 bias-corrected XCO2 and other select fields from the full-physics retrieval aggregated as daily files V7.3, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datacollection/ACOS_L2_Lite_FP_7.3.html.
32. OCO-2 Science Team/Michael Gunson, Annmarie Eldering (2015), OCO-2 Level 2 bias-corrected XCO2 and other select fields from the full-physics retrieval aggregated as daily files, Retrospective processing V7r, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC). https://disc.gsfc.nasa.gov/datacollection/OCO2_L2_Lite_FP_7r.html
33. Schuh, A. E. et al. Quantifying the impact of atmospheric transport uncertainty on $CO_2$ surface flux estimates. *Glob. Biogeochem. Cycles* **33**, 484–500 (2019).
34. Crowell, S. et al. The 2015–2016 Carbon Cycle As Seen from OCO-2 and the Global *In Situ* Network, Atmos. Chem. Phys. Discuss. https://doi.org/10.5194/acp-2019-87, in review (2019).
35. Myneni, R. MCD15A2H MODIS/Terra + Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/modis/mcd15a2h.006 (2015).
36. Didan, K. MOD13C2 MODIS/Terra Vegetation Indices Monthly L3 Global 0.05 Deg CMG V006, NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/modis/mod13c2.006 (2015).
37. Swenson, S. & Wahr, J. Post-processing removal of correlated errors in GRACE data. *Geophys. Res. Lett.* **33**, L08402 (2006).
38. Swenson, S. C. GRACE monthly land water mass grids NETCDF RELEASE 5.0. Ver. 5.0. PO.DAAC, CA, USA. https://doi.org/10.5067/TELND-NC005 (2012).
39. Landerer, F. W. & Swenson, S. C. Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.* **48**, W04531 (2012).

40. Levelt, P. F. et al. The Ozone Monitoring Instrument: overview of 14 years in space. *Atmos. Chem. Phys.* **18**, 5699–5745 (2018).

41. Palmer, P. I. et al. Mapping isoprene emissions over North America using formaldehyde column observations from space. *J. Geophys. Res.* **108**, 4180 (2003).

42. Wolfe, G. M. et al. Formaldehyde production from isoprene oxidation across $NO_x$ regimes. *Atmos. Chem. Phys.* **16**, 2597–2610 (2016).

43. Akagi, S. K. et al. Emission factors for open and domestic biomass burning for use in atmospheric models. *Atmos. Chem. Phys.* **11**, 4039–4072 (2011).

44. Gonzi, S. et al. Biomass burning emission estimates inferred from satellite column measurements of HCHO: sensitivity to co-emitted aerosol and injection height. *Geophys. Res. Lett.* **38**, L14807 (2011).

45. Le Page, Y. et al. Global fire activity patterns (1996–2006) and climatic influence: an analysis using the World Fire Atlas. *Atmos. Chem. Phys.* **8**, 1911–1924 (2008).

46. Somkuti, P. Novel Methods for Atmospheric Carbon Dioxide Retrieval from the JAXA GOSAT and NASA OCO-2 Satellites, PhD Thesis, University of Leicester (2018).

47. Frankenberg, C. et al. New global observations of the terrestrial carbon cycle from GOSAT: patterns of plant fluorescence with gross primary productivity. *Geophys. Res. Lett.* **38**, L17706 (2011).

48. ESA CCI Land Cover Product. https://www.esa-landcover-cci.org/.

49. Van der Werf, G. R. et al. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* **10**, 11707–11735 (2010).

50. Feng, L. et al. Evaluating a 3-D transport model of atmospheric $CO_2$ using ground-based, aircraft, and space-borne data. *Atmos. Chem. Phys.* **11**, 2789–2803 (2011).

51. Feng, L. et al. Estimates of European uptake of $CO_2$ inferred from GOSAT $X_{CO_2}$ retrievals: sensitivity to measurement bias inside and outside Europe. *Atmos. Chem. Phys.* **16**, 1289–1302 (2016).

52. Kawa, S. R., Erickson, D. J., Pawson, S. & Zhu, Z. Global $CO_2$ transport simulations using meteorological data from the NASA data assimilation system. *J. Geophys. Res. -Atmos.* **109**, D18312 (2004).

53. Locatelli, R., Bousquet, P., Saunois, M., Chevallier, F. & Cressot, C. Sensitivity of the recent methane budget to LMDz sub-grid-scale physical parameterizations. *Atmos. Chem. Phys.* **15**, 9765–9780 (2015).

54. Oda, T. & Maksyutov, S. A very high-resolution (1 km × 1 km) global fossil fuel $CO_2$ emission inventory derived using a point source database and satellite observations of nighttime lights. *Atmos. Chem. Phys.* **11**, 543–556 (2011).

55. Oda, T., and Maksyutov, S. ODIAC Fossil Fuel $CO_2$ Emissions Dataset (ODIAC2016), Center for Global Environmental Research, National Institute for Environmental Studies. https://doi.org/10.17595/20170411.001 (2015).

56. Takahashi, T. et al. Climatological mean and decadal change in surface ocean $pCO_2$, and net sea–air $CO_2$ flux over the global oceans. *Deep Sea Res. II* **56**, 554–577 (2009).

57. Olsen, S. C. & Randerson, J. T. Differences between surface and column atmospheric $CO_2$ and implications for carbon cycle research. *J. Geophys. Res. Atmos.* **109**, D02301 (2004).

58. Feng, L., Palmer, P. I., Bösch, H. & Dance, S. Estimating surface $CO_2$ fluxes from space-borne $CO_2$ dry air mole fraction observations using an ensemble Kalman Filter. *Atmos. Chem. Phys.* **9**, 2619–2633 (2009).

59. Saeki, T. & Patra, P. K. Implications of overestimated anthropogenic $CO_2$ emissions on East Asian and global land $CO_2$ flux inversion, *Geosci. Lett.* https://doi.org/10.1186/s40562-017-0074-7 (2017).

60. Dee, D. P. et al. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**, 553–597 (2011).

61. Nasser, R. et al. Improving the temporal and spatial distribution of $CO_2$ emissions from global fossil fuel emission data sets. *J. Geophys. Res.* **118**, 917–933 (2013).

62. Krinner, G. N. et al. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Glob. Biogeochem. Cycles* **19**, DB1015 (2005).

63. Chevallier, F. et al. Inferring $CO_2$ sources and sinks from satellite observations: method and application to TOVS data. *J. Geophys. Res.* **110**, D24309 (2005).

64. Chevallier, F. Description of the $CO_2$ inversion production chain. CAMS deliverable CAMS73_2015SC2_D73.1.5.5_201703_CO2 inversion production chain_v1. http://atmosphere.copernicus.eu/ (2017).

65. Fisher, M. Minimization algorithms for variational data assimilation. In *Proceedings of Seminar on Recent Developments in Numerical Methods for Atmospheric Modelling*. 7–11 September 1998. 364–385. (ECMWF: Reading, UK, 1998).

66. Bosilovich, M. G., Lucchesi, R. & M. Suarez, M. MERRA-2: File Specification. GMAO Office Note No. 9 (Version 1.1). http://gmao.gsfc.nasa.gov/pubs/office_notes (2016).

67. van der Werf, G. R. et al. Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period. *Science* **303**, 73–76 (2004).

68. van der Werf, G. R. et al. Interannual variability of global biomass burning emissions from 1997 to 2004. *Atmos. Chem. Phys.* **6**, 3423–3441 (2006).

69. Gregg, W. W. & Rousseaux, C. S. Decadal trends in global pelagic ocean chlorophyll: A new assessment integrating multiple satellites, in situ data, and models. *J. Geophys. Res. Oceans* **119**, 5921–5933 (2014).

70. Landschützer, P., Gruber, N. & Bakker, D. C. E. A 30-year observation-based global monthly gridded sea surface pCO2 product from 1982 through 2011 http://cdiac.ornl.gov/ftp/oceans/spco2_1982_2011_ETH_SOM-FFN. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee. https://doi.org/10.3334/CDIAC/OTG.SPCO2_1982_2011_ETH_SOM-FFN (2015).

71. Baker, D. F., Bösch, H., Doney, S. C. & Schimel, D. S. Carbon source/sink information provided by column $CO_2$ measurements from the Orbiting Carbon Observatory. *Atmos. Chem. Phys.* **10**, 4145–4165 (2010).

72. Baker, D. F., Doney, S. C. & Schimel, D. S. Variational data assimilation for atmospheric $CO_2$. *Tellus-B* **58**, 359–365 (2006).

## Acknowledgements

## Author contributions

P.I.P. designed the research with D.B., F.C., L.F.; D.B., F.C., and L.F. determined $CO_2$ fluxes from the satellite data; PIP performed the data analysis and wrote the paper, with L.F., D.B., F.C., H.B., and P.S. providing input on the interpretation of the analysis.

## Additional information