Global Atmospheric Carbon Budget: results from an ensemble of atmospheric CO₂ inversions

P. Peylin¹, R. M. Law², K. R. Gurney³, F. Chevallier¹, A. R. Jacobson⁴, T. Maki⁵, Y. Niwa⁵, P. K. Patra⁶, W. Peters⁷, P. J. Rayner¹, ⁸, C. Rödenbeck⁹, X. Zhang³

¹ Laboratoire des Sciences du Climat et de l’Environnement, Gif sur Yvette, France
² Centre for Australian Weather and Climate Research, CSIRO Marine and Atmospheric Research, Aspendale, Australia
³ School of Life Sciences/Global Institute of Sustainability, Arizona State University, Tempe, USA
⁴ University of Colorado and NOAA Earth System Research Laboratory, Boulder, USA
⁵ Meteorological Research Institute, Tsukuba, Japan
⁶ Research Institute for Global Change, JAMSTEC, Yokohama, Japan
⁷ Dept. of Meteorology and Air Quality, Wageningen University, Wageningen, Netherlands
⁸ School of Earth Sciences, University of Melbourne, Parkville, Australia
⁹ Max-Planck-Institute for Biogeochemistry, Jena, Germany

Correspondence to: P. Peylin (peylin@lsce.ipsl.fr)

Supplementary material

Description of inversions
A short description of the different inversions is provided below, and a list of stations for each inversion can be found in the section “Observational constraints used by the participating inversion systems”, or under: https://transcom.lsce.ipsl.fr.

**LSCE_analytical (LSCEa):**

LSCEa corresponds to the results described in Piao et al., 2009, the sensitivity test without Siberian vertical profiles. It is based on a “matrix” formulation (see Peylin et al., 2005). **Fluxes:** solved at the spatial resolution of the transport model and monthly resolution; prior land fluxes taken as the climatology over 1996-2004 from the ORCHIDEE model (Krinner et al., 2005); prior ocean fluxes from Takahashi et al., (2002). Prior land/ocean errors set to 6.0/2.5 Pg C yr\(^{-1}\) globally and spatially distributed according to the Gross Primary Production of ORCHIDEE / the surface area of ocean grid cells; flux error correlations between land/ocean grid-points, following an e-folding length of 1000/2000 km. **Observations:** 73 sites from GLOBALVIEW-CO\(_2\) and CARBOEUROPE EU-project (9 sites); Errors (measurements + model) range between 0.4 ppm for remote stations (South Pole) and 3 ppm for continental sites (Hungaria).

**Prescribed fluxes:** fossil fuel with spatial distribution from Oliver and Berdowski (2001) and annual totals rescaled each year for each country using CDIAC statistics; Biomass burning from van der Werf et al. (2006).

**LSCE_variational (LSCEv):**

LSCEv corresponds to the results described in Chevallier et al., 2010,. It is based on a variational formulation (see Chevallier et al., 2005), with posterior errors provided by a robust Monte Carlo approach. Fluxes: solved at the spatial resolution of the transport model and at weekly resolution; prior land fluxes from the ORCHIDEE model at appropriate date for natural vegetation (Krinner et al. 2005), from the fossil fuel inventory of Oliver and Berdowski (2001) with annual global totals rescaled using CDIAC statistics, from the biomass burning estimates of Randerson et al. (2007); prior ocean fluxes from Takahashi et al., (2009). Prior land/ocean errors set to 3.6/0.9 PgC y\(^{-1}\) globally and spatially distributed according to the heterotrophic respiration of ORCHIDEE / the surface area of ocean grid cells; flux error correlations between land/ocean grid-points, following an e-folding length of 500/1000 km. Observations: 128
sites from a series of global databases; Errors (measurements + model) range between a
few tenths of a ppm for marine stations and up to 6 ppm for continental sites (CBW).

**CarbonTracker US (CT2009):**

CarbonTracker is an ongoing program of the United States National Oceanic and
Atmospheric Administration (NOAA) to publish approximately-annual estimates of CO₂
surface exchange over the globe. The 2009 update of CarbonTracker (CT2009) used here
is a revised version of the system described in Peters et al. (2007) and is fully
documented online at http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2009/.

CarbonTracker uses an ensemble Kalman filter scheme to estimate weekly scaling factors
multiplying prior-model net carbon exchange over 126 land and 30 ocean regions
covering the globe. Flask and quasi-continuous observations from 94 sites of the CO₂
observing networks operated by NOAA, Environment Canada, the Australian
Commonwealth Scientific and Industrial Research Organization (CSIRO), the National
Center for Atmospheric Research (NCAR), and Lawrence Berkeley National Laboratory
are assimilated to produce optimal surface flux estimates. A relatively short assimilation
window of five weeks is used to determine adjustments to surface fluxes. Model-data
mismatch errors assigned to observations range from 0.75 to 7.5 ppm. Atmospheric
transport is simulated with the nested-grid TM5 model described in Krol et al. (2005),
using winds from the operational forecast model of the European Centre for Medium-
Range Weather Forecasts.

CarbonTracker simulates four types of CO₂ exchange with the atmosphere. Fossil-fuel
and biomass burning estimates are imposed without modification, and air-sea exchange
and non-wildfire land exchange are subject to optimization.

Annual country-total fossil fuel emissions for CT2009 were derived from Carbon
Dioxide Information and Analysis Center (CDIAC) estimates of Andres et al. 2009
through 2006, then extrapolated for 2007 and 2008, using relative increases for each fuel
type from the BP Statistical Review of World Energy. These emissions are distributed
spatially according to EDGAR v4 inventory estimates, and an annual cycle is applied
between 30°N and 60°N, as detailed in the online documentation for CT2009. Land
biosphere priors are supplied by the Global Fire Emissions Database (GFED) version 2 of
van der Werf et al. (2006). Air-sea exchange of CO$_2$ is derived from the ocean interior inversions of Jacobson et al. (2007), and includes a trend in anthropogenic CO$_2$ uptake by the world ocean.

Compared to the CT2007 release described in Peters et al (2007), the following significant changes have been made in CT2009: i) Observations from 12 new quasi-continuous stations have been used as assimilation constraints; ii) Seasonality of fossil fuel emissions has been extended to the entire Northern Hemisphere north of 30°N; iii) The air-sea CO$_2$ flux prior is now time-varying and comes from the ocean inversions reported in Jacobson et al. (2007). The number of ocean regions has been increased to 30 from its original 11. The resolution of atmospheric transport in the global domain has been increased to 3˚x2˚ (N. American transport remains at 1˚x1˚).

**CTTracker Europe (CTE2008):**

CarbonTracker Europe is based on the exact same inversion framework as CarbonTracker US described above. It differs in a number of important choices for the inputs, specifically:

- CT Europe uses a set of 23 CarboEurope CO$_2$ mole fraction observations not available in CT US,
- CT US uses CSIRO CO$_2$ mole fraction observations which were not assimilated in CT Europe,
- CT Europe uses a TM5 two-way nested transport grid with highest 1x1 degree resolution over Europe instead of over North America,
- CT Europe uses a different seasonality of European fossil fuel emissions (based on work from the Institut fur Energiewirtschaft und Rationelle Energieanwendung (IER), Stuttgart),
- CT Europe has an increased number of European vegetation types (38 instead of 19 possible types) for which weekly parameters are optimized.

**CCAM and MATCH:**
The CCAM and MATCH inversions use a Bayesian synthesis method and are described in Rayner et al. (2008), except that the time period of the inversions has been extended to 2008 and a slightly different set of observing sites has been used. The CCAM and MATCH inversions set up is identical except for the transport model used (CCAM or MATCH) and the number of regions solved for (CCAM: 94 land, 52 ocean, MATCH: 67 land, 49 ocean). Neither transport model used interannual-varying winds. *Fluxes:* Land fluxes are solved relative to a CASA climatology (Randerson et al., 1997). Most land priors are zero relative to CASA with some non-zero priors representing land-use change. Ocean fluxes are solved relative to the climatology of Takahashi et al. (1999) with prior fluxes of zero. Prior land uncertainties are scaled by NPP while ocean uncertainties are scaled by region area with total uncertainty similar to Baker et al. (2006). *Prescribed fluxes:* fossil emissions follow a spatial distribution which is a linear combination of Andres et al. (1996) representing 1990 and Brenkert (1998) representing 1995, scaled to annual totals from CDIAC. *Observations:* 73 CO₂ records from GLOBALVIEW-CO₂ (2009), used as monthly means, 7 δ¹³CO₂ records from CSIRO (Francey et al., 1996). Data uncertainties range from 0.3-9.2 ppm and vary in time. Larger uncertainties are used for periods with extrapolated data from GLOBALVIEW.

**Jena s96-v3.3 (JENA):**

*Observations:* Data from 52 sites whose records span the whole inversion period (1996-2009). Flask pair values or hourly values, respectively, are used directly at their time of measurement. Hourly data are selected for daytime or nighttime values at certain sites (see Table at the end of the Supplementary material). *Fluxes:* Estimated at the spatial resolution of the transport model and daily time steps, with a-priori spatial and temporal correlations (decaying with distance). Prior fluxes are constant in time, in order not to influence variability of the estimates. *Prescribed fluxes:* Fossil fuel emissions from EDGAR 4.0 (linearly extrapolated after 2005 using BP global totals). *Solution method:* Conjugate Gradients minimization with re-orthonormalization after each iteration. Jena inversion runs are also available for longer time periods (starting 1981 using 15 long-record sites), or using more sites (up to 61, for shorter periods over which all sites
provide data). All results, including regular updates, can be downloaded from “http://www.bgc-jena.mpg.de/~christian.roedenbeck/download-CO2/”.

**TRANSCOM mean (TrC):**

The TransCom mean results are based on the TransCom 3 Level 2 analysis found in Gurney et al. (2008) and Baker et al., 2006, but the observational time series have been extended to 2008 (inclusive). The individual posterior flux results (from 11 transport models) are averaged to generate the multi-model mean. The observational time series spans the 1990 to 2008 time period with a total of 103 observing sites from the GLOBALVIEW-CO$_2$ database. The inversion approach used in the TransCom 3 Level 2 results follows the Bayesian synthesis method (Enting 2002). There are 11 land and 11 ocean basis functions that are roughly sub-continental in size. The four background carbon fluxes consisted of 1990 and 1995 fossil fuel emission fields (Andres et al., 1996; Brenkert, 1998), an annually-balanced, seasonal biosphere exchange (Randerson et al., 1997), and air-sea gas exchange (Takahashi et al., 1999). These fluxes are included in the inversion with a small prior uncertainty so that their magnitude is effectively fixed.

**RIGC TDI-64 (RIGC):**

This Bayesian time-dependent inversion with 64-regions (TDI-64) is developed based on the TransCom level 3 inverse model in order to increase the degrees of freedom for flux estimation (or reduce regional aggregation error). The 11 land and 11 ocean regions are divided into 42 and 22 regions, respectively (detailed sensitivity tests for prior flux and data uncertainties/network are discussed in Patra et al., 2005). By this division, we are able to draw distinction between the east and west or north and south of 10 TransCom land regions, and north and south of the Tropical Asia and all ocean regions.

**Atmospheric CO$_2$** time series from 74 GLOBALVIEW-CO2 sites are used with their corresponding uncertainty derived from climatology of the monthly mean residuals plus 0.3 ppm as a measure of the model representation error. The data uncertainty varies from 0.31 ppm at SPO to 4.6 ppm at HUN and 5.1 ppm at BSC. The NIES/FRCGC transport model (Maksyutov et al., 2008) is driven by interannually varying NCEP reanalysis meteorology. The pre-subtracted fluxes are taken from CASA terrestrial ecosystem model (Randerson et al., 1997) and Takahashi et al. (2009) climatology for oceanic
exchange at monthly time intervals. Fossil fuel emission distributions are taken from EDGAR4.0 and global totals are scaled to CDIAC estimated annual emissions. Prior flux uncertainties are assigned in range of ~0.37 PgC yr$^{-1}$ to ~2.12 PgC yr$^{-1}$ for both land and ocean regions.

**JMA 2010 (JMA):**

JMA inversion method corresponds to the method described in Maki et al. 2010 which is based upon Transcom 3 IAV inversion set up (Baker et al. 2006) with real observation data (WDCGG) and interannual varying wind (JRA25). The analysis period is extended and some modifications are done. **Fluxes:** solved at the spatial resolution of 22 regions and monthly resolution; prior land fluxes taken as the climatology from CASA model; prior ocean fluxes from Takahashi et al., (2002). Prior land/ocean errors set to Transcom 3 IAV uncertainties; flux error correlations are set to zero. **Observations:** 146 sites from WDCGG monthly mean CO$_2$ concentrations after site selection by mismatch between observation and inversion; Errors (measurements + model) range between 0.3 ppm for remote stations (South Pole) and 5 ppm for continental sites. **Prescribed fluxes:** fossil fuel with spatial distribution from Andres et al. (1996) and Brenkert (1998) annual totals rescaled each year for each country using CDIAC statistics.

**NICAM-TM (NICAM):**

NICAM-TM inversion system is described by Niwa et al. (2012). While Niwa et al. (2012) extensively used aircraft measurements from CONTRAIL this inversion used only surface measurements and limited aircraft measurements. The inversion method and setup are similar to those of TRANSCOM. **Fluxes:** the spatial number of fluxes solved by the inversion is 29 and 11 respectively for land and ocean. The 29 land regions were obtained by dividing the 11 regions of TRANSCOM (slightly different from 31 regions of Niwa et al. (2012)). The ocean flux partitioning is the same as TRANSCOM. The prior land flux is taken from the climatology flux of CASA (Randerson et al. 1997); the prior ocean flux is from Takahashi et al., (2009). Prior land flux errors are given by redistributing those for the 11 regions of TRANSCOM into the 29 regions according to NPP distributions, whereas prior ocean flux errors are the same as those of TRANSCOM. There is no error correlation for the prior fluxes. **Observations:** 94 sites from
GLOBALVIEW-CO$_2$ and 9 sites from Siberian aircraft data of NIES; Errors (measurements + model) range between 0.3 ppm for remote stations and 6.6 ppm for continental sites (LEF) (monthly mean). \textit{Prescribed fluxes}: fossil fuel with spatial distribution from EDGAR version 4.1 and annual totals rescaled each year for each country using CDIAC statistics.
**Additional figures**

This appendix provides additional figures showing the estimated and prior aggregated carbon fluxes (Figures S1 to S6) as well as the region boundaries used to aggregate the fluxes (Figure S7) and the spatial flux distribution (Figure S8).

Figure S1: Annual mean posterior flux of the individual participating inversions for natural global total carbon exchange without fossil correction.
Figure S2: Annual mean posterior natural land flux estimate of the individual participating inversions. Shown here are a) North (>25N) b) Tropics (25S<<25N), c) South (<25S) d) North America, e) Europe, f) North Asia, g) South America, h) Africa, i) South Asia.
Figure S3: Same as Fig. S2 but for the Prior land fluxes.
Figure S4: Annual mean posterior natural ocean flux estimate of the individual participating inversion. Shown here are a) North (>25N) b) Tropics (25S<<25N), c) South (<25S) d) North Pacific, e) North Atlantic, f) Tropical Indian ocean, g) Tropical Pacific, h) Tropical Atlantic, i) Sub tropical ocean.
Figure S5: Same as Fig. S4 but for the Prior ocean fluxes.
Figure S6: Mean seasonal cycle for the prior fluxes from most participating models for selected regions

Figure S7: Region boundaries used for the aggregated fluxes: a) latitudinal breakdown (north, tropics, south) for land and ocean; b) continental breakdown used for north America, Europe, north Asia, south America, Africa, tropical Asia, Australia, north Pacific, north Atlantic, tropical pacific, tropical Atlantic, Indian ocean, sub-tropical ocean.
Figure S8: Spatial distribution of the annual natural fluxes (without fossil correction) for 2003 for the participating inversions.
Observational constraints used by the participating inversion systems

Sites are listed alphabetically, in general using the site codes of GLOBALVIEW (www.esrl.noaa.gov/gmd/ccgg/globalview/co2/co2_observations.html). Many site locations are listed more than once, since multiple CO₂ records are available for many sites (either collected by different labs, or representing separate flask and in-situ records). Some inversions chose only to use the most complete record at a given location while others include all available records. It is unlikely that one choice is better than the other. There are differences in calibration etc between laboratories. These are not accounted for in inversions, but their impact is unlikely to be significant compared to other transport and representation uncertainties in modelling any given site (Rödenbeck et al., 2006).

Type of observed CO₂ used for each site: There are a variety of ways the CO₂ data from any given site has been used, depending in part on whether flask or in-situ data are available; some inversions use monthly mean data while others use the data at the appropriate sampling time. Various degrees of selection have been applied to the data. An indication of how the data have been used is given in the table through a series of codes.

<table>
<thead>
<tr>
<th>Code</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal resolution</strong></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>Monthly mean data used</td>
</tr>
<tr>
<td>D</td>
<td>Daily mean data used</td>
</tr>
<tr>
<td>H4</td>
<td>Four hour mean data used</td>
</tr>
<tr>
<td>H4 (hrs)</td>
<td>Four hour mean data used only for the hours indicated (UT)</td>
</tr>
<tr>
<td>H</td>
<td>Hourly data used</td>
</tr>
<tr>
<td>H (hrs)</td>
<td>Hourly data used only for the hours indicated (UT)</td>
</tr>
<tr>
<td>F4</td>
<td>Flask samples are used as a 4 hour average around sampling time</td>
</tr>
<tr>
<td>F</td>
<td>Flasks samples are used at the sampling time</td>
</tr>
<tr>
<td><strong>Data selection</strong></td>
<td></td>
</tr>
<tr>
<td>GV</td>
<td>Globalview CO₂ used. Globalview CO₂ is derived from a fitted curve to CO₂ observations</td>
</tr>
</tbody>
</table>
and is intended to represent baseline conditions

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>GV E</td>
<td>The use of Globalview CO₂ data includes extrapolated data. This fills in missing data by applying a mean seasonal offset for the site from marine boundary layer CO₂ concentration. Many inversions give periods of extrapolated data less weight than periods with observations.</td>
</tr>
<tr>
<td></td>
<td>* Consecutive hours that differ by greater than 1 ppm are removed</td>
</tr>
<tr>
<td></td>
<td>o Outliers (mismatch between observations and model &gt; 3 sigma) are removed</td>
</tr>
<tr>
<td>JMA</td>
<td>Data are removed when inconsistent with the inversion through an iterative procedure (Maki et al., 2010)</td>
</tr>
</tbody>
</table>

Site locations may be represented by model output interpolated to the site location or by the nearest model grid-cell. For coastal sites, the nearest ocean grid-cell if often chosen as being more representative of the baseline air that is usually sampled by flask records at coastal sites.

Some inversions include ship data. This may be used at the actual location and time of sampling (as in the CarbonTracker inversions) or it may be binned into latitude and/or longitude bins as in the GLOBALVIEW POC* records and the JMA use of JMA ship data.