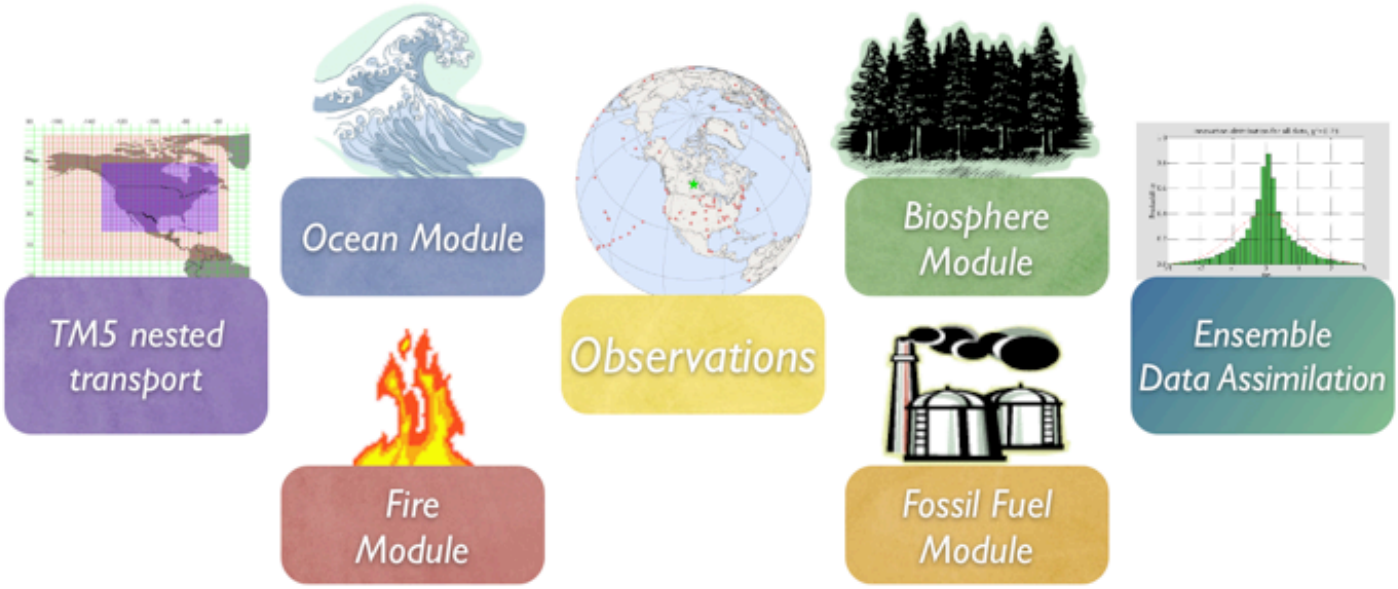
Documentation



To learn more about a CarbonTracker component, click on one of the above images.
Or download the full PDF version for convenience.

Oceans Module [goto top]

1. Introduction

The oceans play an important role in the Earth's carbon cycle. They are the largest long-term sink for carbon and have an enormous capacity to store and redistribute CO_2 within the system. A large fraction of the CO_2 burned by fossil fuels and land-use change since the Industrial Revolution has been absorbed by the world's oceans, leading to a measurable acidification (dissolving CO_2 in sea water increases the number of H^+ ions and thereby the pH-value). Although the oceans as a whole have been a steady carbon sink, CO_2 an also come out of the oceans depending on local temperatures, biological activity, wind speeds, and ocean currents. Such effects have to be included in CarbonTracker to monitor the behavior of the ocean sink, and to understand variability in other carbon cycle components (i.e., the biosphere).

2. Detailed Description

Ocean exchange of CO_2 in our system is computed using climatological air-sea differences in partial pressure of CO_2 combined with 3-hourly wind speed and barometric pressure from the atmospheric transport model.

Seawater pCO_2 is provided by Takahashi et al. [2002]. These are monthly climatological estimates of CO_2 partial pressure in surface waters, upsampled from the native 4x5 degree resolution to a 1x1 degree grid. This climatology is formed from about 940,000 discrete measurements of surface water pCO_2. We also use the atmospheric pCO_2 from the Takahashi et al. [2002] climatology, instead of the modeled atmospheric pCO_2. Available observations suggest that the air-sea CO_2 disequilibrium remains fairly constant in the face of rising atmospheric concentrations [Takahashi et al., 2006]. Atmospheric pCO_2 is, however, modulated by the surface barometric pressure in the atmospheric transport model following the formulation of Kettle and Merchant [2005].

The gas transfer coefficient (k) is parameterized following the quadratic wind speed formulation of Wanninkhof [1992] for instantaneous winds. Gas exchange is computed every 3 hours using the European Centre for Medium-Range Weather Forecasts (ECMWF) forecast meteorology of the atmospheric transport model. The introduction of high resolution wind speed and pressure variability on the ocean fluxes follows the methods of Bates and Merlivat [2001] and Kettle and Merchant [2005]. The latter study found a 7% decrease in annual mean ocean sink when covariations of wind and pressure are taken into account.

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

3. Further Reading

- WikiPedia
- Takahashi et al., data
- Takahashi et al., live data access
- Kettle and Merchant, paper
- Jacobson, paper

Biosphere Module [goto top]

1. Introduction

The biospheric component of the carbon cycle consists of all the carbon stored in 'biomass' around us. This includes trees, shrubs, grasses, carbon within soils, dead wood, and leaf litter. Such reservoirs of carbon can exchange CO_2 with the atmosphere. Exchange starts when plants take up CO_2 during their growing season through the

process called photosynthesis (uptake). Most of this carbon is released back to the atmosphere throughout the year through a process called respiration (release). This includes both the decay of dead wood and litter and the metabolic respiration of living plants. Of course, plants can also return carbon to the atmosphere when they burn, as described here. Even though the yearly sum of uptake and release of carbon amounts to a relatively small number (a few petagrams (one Pg=10^{15} g)) of carbon per year, the flow of carbon each way is as large as 120 Pg each year. This is why the net result of these flows needs to be monitored in a system such as ours. It is also the reason we need a good physical description (model) of these flows of carbon. After all, from the atmospheric measurements we can only see the small net sum of the large two-way streams (gross fluxes). Information on what the biospheric fluxes are doing in each season, and in every location on Earth is derived from a specialized biosphere model, and fed into our system as a first guess, to be refined by our assimilation procedure.

2. Detailed Description

The biosphere model currently used in CarbonTracker is the Carnegie-Ames Stanford Approach (CASA) biogeochemical model. This model calculates global carbon fluxes using input from weather models to drive biophysical processes, as well as satellite observed Normalized Difference Vegetation Index (NDVI) to track plant phenology. The version of CASA model output used so far was driven by year specific weather and satellite observations, and including the effects of fires on photosynthesis and respiration (see van der Werf et al., [2006] and Giglio et al., [2006]). This simulation gives 1x1 degree global fluxes on a monthly time resolution.

Net Ecosystem Exchange (NEE) is re-created from the monthly mean CASA Net Primary Production (NPP) and ecosystem respiration (R_E). Higher frequency variations (diurnal, synoptic) are added to Gross Primary Production (GPP=2*NPP) and R_E(=NEE-GPP) fluxes every 3 hours using a simple temperature Q_{10} relationship assuming a global Q_{10} value of 1.5 for respiration, and a linear scaling of photosynthesis with solar radiation. The procedure is very similar, but **NOT** identical to the procedure in Olsen and Randerson [2004] and based on ECMWF analyzed meteorology. Note that the introduction of 3-hourly variability conserves the monthly mean NEE from the CASA model. Instantaneous NEE for each 3-hour interval is thus created as:

$$NEE(t) = GPP(I, t) + R_E(T, t)$$

$$GPP(t) = I(t) * (\sum(GPP) / \sum(I))$$

$$R_E(t) = Q_{10}(t) * (\sum(R_E) / \sum(Q_{10}))$$

$$Q_{10}(t) = 1.5^{((T_{2m}-T_0) / 10.0)}$$

where T=2 meter temperature, I=incoming solar radiation, t=time, and summations are done over one month in time, per gridbox. The instantaneous fluxes yielded realistic diurnal cycles when used in the TransCom Continuous experiment.

The current CarbonTracker v7.0 release was based on the CASA runs for the GFED2 project to estimate fire emissions. We found a significantly better match to observations when using this output compared to the fluxes from a neutral biosphere simulation. Due to the inclusion of fires, inter-annual variability in weather and NDVI, the fluxes for North America start with a small net flux even when no assimilation is done. This flux ranges from 0.05 PgC/yr of release, to 0.15 PgC/yr of uptake.

3. Further Reading

- CASA with fires model overview
- CASA results from Jim Randerson
- GFED2 results from Guido van der Werf, Jim Randerson, and colleagues
- Olsen and Randerson, paper
- Giglio et al., 2006 paper
- van der Werf et al., 2006 paper

Fires Module [goto top]

1. Introduction

Fires are an important part of the carbon cycle and have been so for many millennia. Even before human civilization began to use fires to clear land for agricultural purposes, most ecosystems were subject to natural wildfires that would rejuvenate old forests and bring important minerals to the soils. When fires consume part of the landscape in either controlled or natural burning, carbon dioxide (amongst many other gases and aerosols) is released in large quantities. Each year, more than 2 PgC is estimated to be released from the biosphere into the atmosphere through fires, mostly in the tropics. Nowadays, a large fraction of these fires is started by humans, and mostly intentionally to clear land for agriculture, or to re-fertilize soils before a new growing season. This important component of the carbon cycle is monitored mostly from space, while sophisticated 'biomass burning' models are used to estimate the emissions of CO_2. Such estimates are then used in CarbonTracker to prescribe the emissions, without further refinement by our measurements.

2. Detailed Description

The fire module currently used in CarbonTracker is based on the Global Emissions Fire Database version 2 (GFED2), which is derived from the CASA biogeochemical model as described here. The dataset consists of 1x1 degree gridded monthly burned area, fuel

loads, combustion completeness, and fire emissions (Carbon, CO_2, CO, CH_4, NMHC, H_2, NO_x, N_2O, PM2.5, TPM, TC, OC, BC) for the time period January 1997 - December 2005, of which we currently only use CO_2. The GFED2 burned area is based on MODIS satellite observations of fire counts. These, together with detailed vegetation cover information and a set of vegetation specific scaling factors, allow predictions of burned area over the time span that active fire counts from MODIS are available. The relationship between fire counts and burned area is derived, for the specific vegetation types, from a 'calibration' subset of 500m resolution burned area from TERRA-MODIS in the period 2001-2004.

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

GFED2 products were successfully used in recent studies of CH_4, CO_2, CO, and other trace gases.

3. Further Reading

- CASA with fires model overview
- CASA results from Jim Randerson
- GFED2 results from Guido van der Werf, Jim Randerson, and colleagues
- Giglio et al., 2006 paper
- van der Werf et al., 2006 paper

Observations [goto top]

1. Introduction

The observations of CO_2 mole fraction by NOAA ESRL and partner laboratories are at the heart of CarbonTracker. They inform us on changes in the carbon cycle, whether they are regular (such as the seasonal growth and decay of leaves and trees), or irregular (such as the release of tons of carbon by a wildfire). The results in CarbonTracker depend directly on the quality and amount of observations available, and the degree of detail at which we can monitor the carbon cycle reliably increases strongly with the density of our observing network.

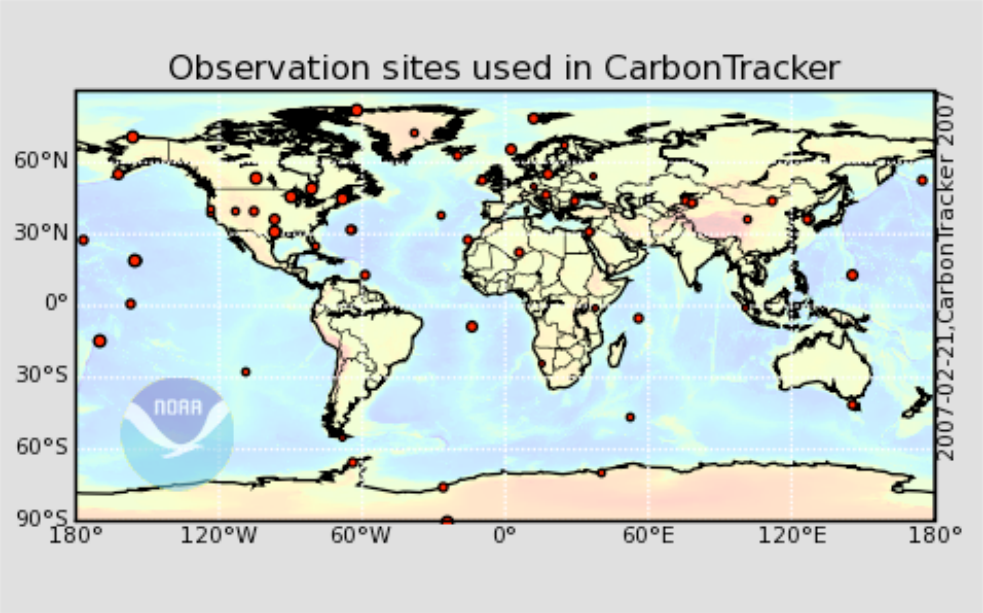
2. Detailed Description

This study uses all analyzed air samples taken at the surface from the NOAA ESRL Cooperative Air Sampling Network available for each year studied, except those flagged for analysis or sampling problems. Assimilated data includes values that are flagged as not representative of typical background conditions. The composition of the network

thus varies per week depending on successful sampling and analysis, and each sites' sampling frequency. In addition, we use continuous CO_2 time series from five towers: (1) the 396m level of the WLEF tower in Wisconsin, (2) the 107m level of the AMT tower in Argyle, Maine, (3) the 251m level of the KWKT tower in Texas, (4) the 40m level of the tower in Fraserdale, Canada operated by the Meteorological Service Canada (MSC), and (5) the 23m level of the tower at Old Black Spruce, Canada operated by MSC. Other in-situ continuous CO_2 time series used are from the NOAA ESRL observatories at Barrow, Mauna Loa, Samoa, and South Pole, and the continuous analyzer at Alert, Canada, operated by MSC. Note that all of these observations are calibrated against the same world CO_2 standard (WMO-2005). Also, note that aircraft observations from the NOAA ESRL program were NOT assimilated, but used for independent assessment of the results.

At the continuous sampling sites, we construct one daytime average (12:00-16:00 Local Standard Time) mole fraction for each day from the time series, recognizing that our atmospheric transport model not always captures the continental nighttime stability regime while daytime well-mixed conditions are better matched. This approach is partly based on analysis of TransCom Continuous (TC) results [Law et al., 2007]. Moreover, observations at sub-daily time scales are likely to be strongly correlated and therefore add relatively little independent information to our results. Also based on TC continuous simulations, we decided to move a set of coastal sites by one degree into the ocean to force the model sample to be more representative of the actual site conditions. These sites are labeled for reference in the complete table of sites used in CarbonTracker.

We apply a further selection criterion during the assimilation to exclude non-Marine Boundary Layer (MBL) observations that are very poorly forecasted in our framework. We use the so-called model-data mismatch in this process, which is the random error prescribed to each observation to account for measurement errors as well as modeling errors of that observation. We interpret an observed-minus-forecasted (OmF) mole fraction that exceeds 3 times the prescribed model-data mismatch as an indicator that our modeling framework fails. This can happen for instance when an air sample is representative of local exchange not captured well by our 1x1 degree fluxes, when local meteorological conditions are not captured by our offline transport fields, but also when large-scale CO_2 exchange is suddenly changed (e.g. fires, pests, droughts) to an extent that can not be accommodated by our flux modules. This last situation would imply an important change in the carbon cycle and has to be recognized by the researchers when analyzing the results. In accordance with the 3-sigma rejection criterion, ~2% of the observations are discarded through this mechanism in our assimilations. Table 1 gives a summary of the sites used and the assimilation performance.



Code	Name	Lat, Lon, Elev	Lab	N used	N flagged	mismatch	Inn X2
alt_06C0	Alert, Nunavut, Canada	82 27'N, 62 31'W, 200.0m	MSC	1737	0	2.50	0.20
amt_01C3	Argyle, Maine, United States	45 2'N, 68 41'W, 50.0m	NOAA	654	52	3.00	0.57
brw_01C0	Barrow, Alaska, United States	71 19'N, 156 36'W, 11.0m	NOAA	1765	2	2.50	0.27
frd_06C0	Fraserdale, Canada	49 53'N, 81 34'W, 210.0m	MSC	1711	40	3.00	0.42
mlo_01C0	Mauna Loa, Hawaii, United States	19 32'N, 155 35'W, 3397.0m	NOAA	1197	0	0.75	0.89
obs_06C0	Old Black Spruce, Saskatchewan, Canada	53 59'N, 105 7'W, 629.0m	MSC	687	4	3.00	0.55
smo_01C0	Tutuila, American Samoa	14 14'S, 170 34'W, 42.0m	NOAA	1965	0	0.75	0.66
spo_01C0	South Pole, Antarctica, United States	89 59'S, 24 48'W, 2810.0m	NOAA	2074	0	0.75	0.45
wkt_01C3	Moody, Texas, United States	31 19'N, 97 20'W, 251.0m	NOAA	544	26	3.00	0.54
lef_01C3	Park Falls, Wisconsin, United States	45 56'N, 90 16'W, 472.0m	NOAA	1532	31	3.00	0.54
alt_01D0	Alert, Nunavut, Canada	82 27'N, 62 31'W, 200.0m	NOAA	279	0	1.50	0.45
asc_01D0	Ascension Island, United Kingdom	7 55'S, 14 25'W, 54.0m	NOAA	499	0	0.75	1.31
ask_01D0	Assekrem, Algeria	23 11'N, 5 25'E, 2728.0m	NOAA	258	0	1.50	0.39
azr_01D0	Terceira Island, Azores, Portugal	38 46'N, 27 23'W, 40.0m	NOAA	189	4	1.50	0.91

bal_01D0	Baltic Sea, Poland	55 21'N, 17 13'E, 3.0m	NOAA	447	1	7.50	0.45
bkt_01D0	Bukit Kototabang, Indonesia	0 12'S, 100 19'E, 864.5m	NOAA	74	0	7.50	0.55
bme_01D0	St. Davids Head, Bermuda, United Kingdom	32 22'N, 64 39'W, 30.0m	NOAA	189	9	1.50	1.26
bmw_01D0	Tudor Hill, Bermuda, United Kingdom	32 16'N, 64 53'W, 30.0m	NOAA	203	3	1.50	1.00
brw_01D0	Barrow, Alaska, United States	71 19'N, 156 36'W, 11.0m	NOAA	269	6	1.50	0.65
bsc_01D0	Black Sea, Constanta, Romania	44 10'N, 28 41'E, 3.0m	NOAA	247	1	7.50	0.75
cba_01D0	Cold Bay, Alaska, United States	55 12'N, 162 43'W, 25.0m	NOAA	438	22	1.50	1.18
cgo_01D0	Cape Grim, Tasmania, Australia	40 41'S, 144 41'E, 94.0m	NOAA	245	0	1.50	0.20
chr_01D0	Christmas Island, Republic of Kiribati	1 42'N, 157 10'W, 3.0m	NOAA	223	0	0.75	1.19
crz_01D0	Crozet Island, France	46 27'S, 51 51'E, 120.0m	NOAA	180	0	0.75	0.96
eic_01D0	Easter Island, Chile	27 9'S, 109 27'W, 50.0m	NOAA	129	0	7.50	0.04
gmi_01D0	Mariana Islands, Guam	13 26'N, 144 47'E, 1.0m	NOAA	463	0	1.50	0.47
hba_01D0	Halley Station, Antarctica, United Kingdom	75 35'S, 26 30'W, 30.0m	NOAA	274	0	0.75	0.89
hun_01D0	Hegyhatsal, Hungary	46 57'N, 16 39'E, 248.0m	NOAA	274	3	7.50	0.44
ice_01D0	Storhofdi, Vestmannaeyjar, Iceland	63 20'N, 20 17'W, 118.0m	NOAA	258	2	1.50	0.52
izo_01D0	Tenerife, Canary Islands, Spain	28 18'N, 16 29'W, 2360.0m	NOAA	209	2	1.50	0.92
key_01D0	Key Biscayne, Florida, United States	25 40'N, 80 12'W, 3.0m	NOAA	196	0	2.50	0.31
kum_01D0	Cape Kumukahi, Hawaii, United States	19 31'N, 154 49'W, 3.0m	NOAA	266	0	1.50	0.49
kzd_01D0	Sary Taukum, Kazakhstan	44 27'N, 75 34'E, 412.0m	NOAA	271	55	2.50	0.65
kzm_01D0	Plateau Assy, Kazakhstan	43 15'N, 77 53'E, 2519.0m	NOAA	238	1	2.50	1.13
mhd_01D0	Mace Head, County Galway, Ireland	53 20'N, 9 54'W, 25.0m	NOAA	224	0	2.50	0.23
mid_01D0	Sand Island, Midway, United States	28 13'N, 177 23'W, 3.7m	NOAA	265	0	1.50	0.62
mkn_01D0	Mt. Kenya, Kenya	0 3'S, 37 18'E,	NOAA	57	0	2.50	1.00

mkn_01D0	Mt. Kenya, Kenya	3897.0m	NOAA	37	0	2.50	1.00
mlo_01D0	Mauna Loa, Hawaii, United States	19 32'N, 155 35'W, 3397.0m	NOAA	310	0	1.50	0.28
nmb_01D0	Gobabeb, Namibia	23 35'S, 15 2'E, 456.0m	NOAA	16	0	2.50	0.17
nwr_01D0	Niwot Ridge, Colorado, United States	40 3'N, 105 35'W, 3523.0m	NOAA	264	2	1.50	0.74
obn_01D0	Obninsk, Russia	55 7'N, 36 36'E, 183.0m	NOAA	72	0	7.50	0.42
oxk_01D0	Ochsenkopf, Germany	50 4'N, 11 48'E, 1193.0m	NOAA	19	1	2.50	1.07
pal_01D0	Pallas-Sammaltunturi, GAW Station, Finland	67 58'N, 24 7'E, 560.0m	NOAA	145	1	2.50	0.82
poc_01D0	Pacific Ocean, N/A	99 59'S, 999 59'W, 10.0m	NOAA	896	1	7.50	0.01
psa_01D0	Palmer Station, Antarctica, United States	64 55'S, 64 0'W, 10.0m	NOAA	276	0	0.75	1.16
pta_01D0	Point Arena, California, United States	38 57'N, 123 44'W, 17.0m	NOAA	180	0	7.50	0.34
rpb_01D0	Ragged Point, Barbados	13 10'N, 59 26'W, 45.0m	NOAA	267	0	1.50	0.72
sey_01D0	Mahe Island, Seychelles	4 40'S, 55 10'E, 3.0m	NOAA	256	0	0.75	1.26
sgp_01D0	Southern Great Plains, Oklahoma, United States	36 48'N, 97 30'W, 314.0m	NOAA	369	29	2.50	0.49
shm_01D0	Shemya Island, Alaska, United States	52 43'N, 174 6'E, 40.0m	NOAA	234	1	2.50	0.82
smo_01D0	Tutuila, American Samoa	14 14'S, 170 34'W, 42.0m	NOAA	315	0	1.50	0.18
spo_01D0	South Pole, Antarctica, United States	89 59'S, 24 48'W, 2810.0m	NOAA	291	0	1.50	0.09
stm_01D0	Ocean Station M, Norway	66 0'N, 2 0'E, 0.0m	NOAA	504	5	1.50	0.76
sum_01D0	Summit, Greenland	72 35'N, 38 29'W, 3238.0m	NOAA	178	0	1.50	0.46
syo_01D0	Syowa Station, Antarctica, Japan	69 0'S, 39 35'E, 11.0m	NOAA	139	0	0.75	1.34
tap_01D0	Tae-ahn Peninsula, Republic of Korea	36 44'N, 126 8'E, 20.0m	NOAA	202	0	7.50	0.38
tdf_01D0	Tierra Del Fuego, Ushuaia, Argentina	54 52'S, 68 29'W, 20.0m	NOAA	74	0	0.75	0.67
thd_01D0	Trinidad Head, California, United States	41 3'N, 124 9'W, 107.0m	NOAA	134	0	7.50	0.44
uta_01D0	Wendover, Utah, United States	39 54'N, 113 43'W, 1320.0m	NOAA	250	55	2.50	0.32
uum_01D0	Ulaan Uul, Mongolia	44 27'N, 111 6'E, 914.0m	NOAA	273	5	2.50	0.67

wis_01D0	Sede Boker, Negev Desert, Israel	31 8'N, 34 53'E, 400.0m	NOAA	296	1	2.50	0.79
wlg_01D0	Mt. Waliguan, Peoples Republic of China	36 17'N, 100 54'E, 3810.0m	NOAA	166	6	1.50	1.16
zep_01D0	Ny-Alesund, Svalbard, Norway and Sweden	78 54'N, 11 53'E, 475.0m	NOAA	355	2	1.50	0.70
all	Total			27675	373	0.00	0.55

3. Further Reading

- ESRL Carbon Cycle Program
- WMO/GAW Report No. 168, 2006

Fossil Fuel Module [goto top]

1. Introduction

Human beings first influenced the carbon cycle through land-use change. Early humans used fire to control animals and later cleared forest for agriculture. Over the last two centuries, following the industrial and technical revolutions and the world population increase, fossil fuel combustion has become the largest anthropogenic source of CO_2. Coal, oil and natural gas combustion indeed are the most common energy source in both developed and developing countries. Various sectors of the economy rely on fossil fuel combustion: power generation, transportation, residential/commercial building heating, and industrial processes. In 2004, the world emissions of CO_2 from fossil fuel burning, cement manufacturing, and flaring reached 7.3 PgC (one Pg=10^{15} g) [CDIAC]. This represents a 26% increase compared to 1990. The North American (U.S.A, Canada, and Mexico) flux of fossil fuel CO_2 to the atmosphere was 1.8 PgC in 2002, representing 26% of the global total. The International Energy Outlook projects that the global total source will reach 9.2 PgC in 2015 and 11.9 PgC in 2030 [DOE].

2. Detailed Description

The fossil fuel emission inventory currently used in CarbonTracker is derived from independent global total and spatially-resolved inventories. Annual global total fossil fuel CO_2 emissions are from Marland et al. [2006] which extend through 2003 and are then linearly extrapolated through 2005. Fluxes are then spatially distributed according to the EDGAR inventories [EDGAR, Olivier and Berdowski, 2001]. The 1x1 EDGAR maps from 1995 and 2000 are used to extrapolate the emission pattern through 2005. Additionally, a seasonal cycle based on the Blasing et al. [2005] analysis for the United States, which has ~20% higher emissions in winter than in summer, is imposed on the North American emissions. The uncertainty attached to the total source is of the order of 15%. This source is not optimized in the current CarbonTracker system as we do not believe our current network can constrain this source separately from the others.

Although the contribution of CO_2 from fossil fuel to the observed CO_2 mole fraction is considered known, extra model error is included in the system to represent the random errors in fossil fuels.

3. Further Reading

- Carbon Dioxide Information Analysis Center (CDIAC)
- Energy Information Administration (EIA)
- EDGAR Database
- CDIAC (Marland et al.) Annual Global fluxes
- CDIAC (Blasing et al.) Monthly USA fluxes

TM5 Nested Transport [goto top]

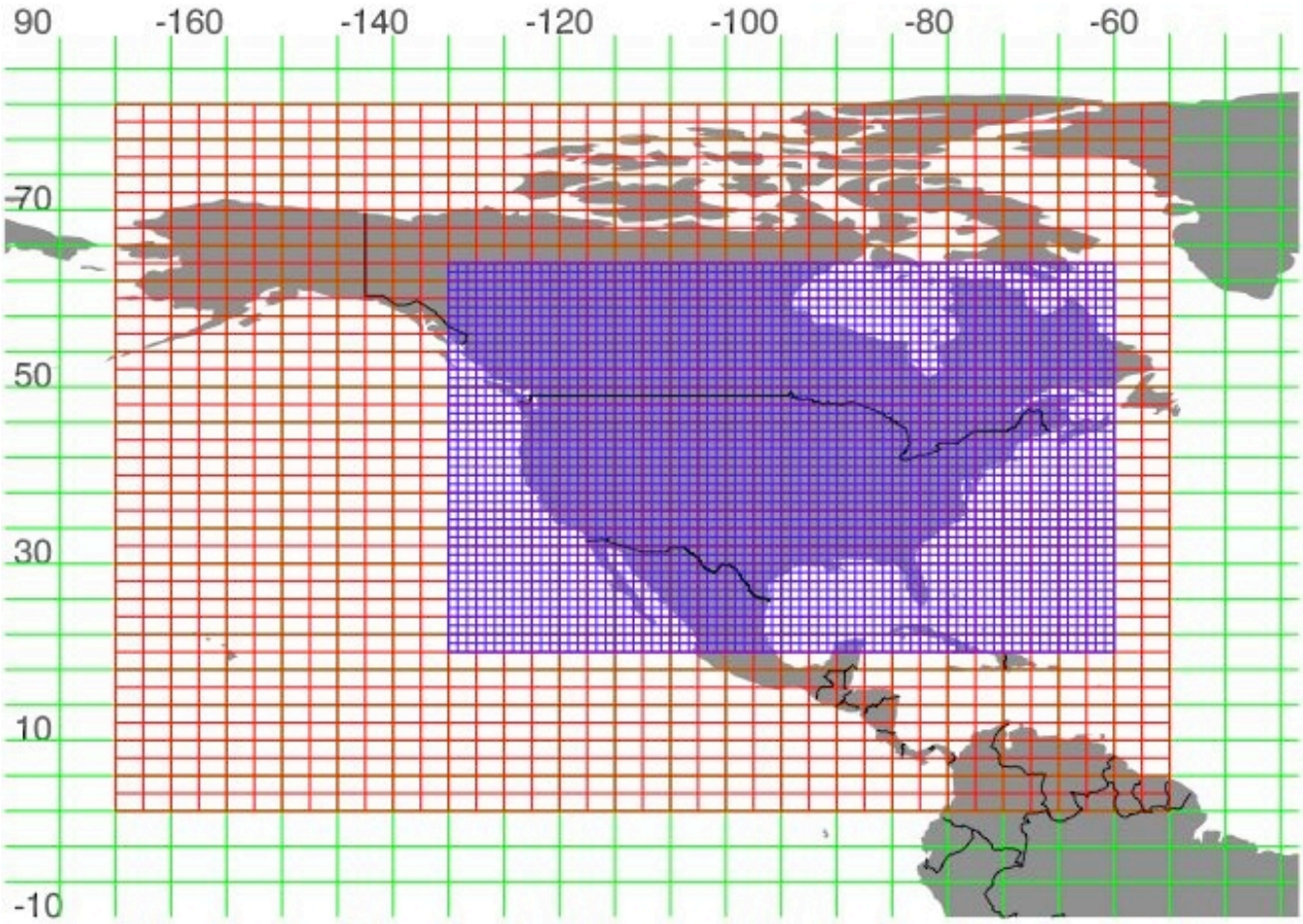
1. Introduction

The link between observations of CO_2 in the atmosphere and the exchange of CO_2 at the Earth's surface is formed by transport in the atmosphere: storm systems, cloud complexes, and weather of all sorts that transports CO_2 around the world and causes local events (fires, forest growth, ocean upwelling) to have impacts on much larger scales. (However, unlike in some climate models, the meteorology we use is fixed and we do not allow the carbon fluxes we estimate to feed back into our transport simulation.) To simulate the winds and the weather, CarbonTracker uses sophisticated numerical models that are driven by the daily weather forecasts from the specialized meteorological centers of the world to transport CO_2 across large distances. Since CO_2 does not decay or react in the lower atmosphere, the influence of emissions and uptake in for instance North America and Europe are ultimately seen in our measurements even on the South Pole! Getting the transport of CO_2 just right is an enormous challenge, and costs us almost 90% of the computer resources for CarbonTracker. The development of the atmospheric transport model, in our case the Transport Model 5 (TM5), is shared among many scientific groups with different expertise. In addition to CarbonTracker, TM5 is also used to make air-quality forecast, study the dispersion of aerosols in the tropics, track biomass burning plumes, and to predict what pollution levels future generations might have to deal with.

2. Detailed Description

TM5 is a global model with the option for two-way nested grids; arbitrary regions with high-resolution grids can be nested in a much coarser grid spanning the global domain. Thus, transport simulations can be performed with a regional focus, without the need for boundary conditions from other models. This approach allows measurements outside the domain of interest to constrain regional fluxes in the data assimilation, and ensures that regional estimates are consistent with global constraints. TM5 is based on its

predecessor TM3, with improvements in the advection scheme, vertical diffusion parameterization, and meteorological preprocessing of the wind fields (Krol et al., 2005). The model is developed and maintained jointly by the Institute for Marine and Atmospheric Research Utrecht (IMAU, The Netherlands), the Joint Research Centre (JRC, Italy), the Royal Netherlands Meteorological Institute (KNMI, The Netherlands), and NOAA ESRL (USA). In CarbonTracker, TM5 simulates the separate processes of advection, convection (deep and shallow), and vertical diffusion in the planetary boundary layer and free troposphere.



TM5 is driven by meteorological data from the European Center for Medium range Weather Forecast (ECMWF) model that currently runs with ~25 km horizontal resolution and 91 layers in the vertical (T799L91) and can also use meteorological data from the National Center for Environmental Prediction (NCEP) forecast. The preprocessing of this data for use in TM5 is done routinely for the global domain (3x2 degrees), as well as for several nested 1x1 degree domains (North America, Europe, Asia). In CarbonTracker, TM5 is run at a global 6x4 degrees resolution with nested regions over North America (3x2 degrees) and the United States (1x1 degree) similar to the set-up in Peters et al., [2004, 2005]. The grid over North America is shown in the figure. TM5 runs at an external time step of three hours, but due to the symmetrical operator splitting and the refined resolution in nested grids, processes at the finest scale are repeated every 10

minutes. The vertical resolution of TM5 in CarbonTracker is 25 hybrid sigma-pressure levels, unevenly spaced with a larger density near the surface. Approximate heights of the mid-levels (in meters, with a surface pressure of 1012 hPa) are:

Level	Height (m)	Level	Height (m)
1	34.5	14	9076.6
2	111.9	15	10533.3
3	256.9	16	12108.3
4	490.4	17	13874.2
5	826.4	18	15860.1
6	1274.1	19	18093.2
7	1839.0	20	20590.0
8	2524.0	21	24247.3
9	3329.9	22	29859.6
10	4255.6	23	35695.0
11	5298.5	24	42551.5
12	6453.8	25	80000.0
13	7715.4		

3. Further Reading

- TM5 technical documentation page (including nightly export of the most recent TM5 manual)
- The TM5 model homepage
- ECMWF forecast model technical documentation
- The NCEP reanalysis meteo data
- Peters et al., 2004, JGR paper on transport in TM5
- Krol et al., 2005, ACP overview paper of the TM5 model

Ensemble Data Assimilation [goto top]

1. Introduction

Data assimilation is the name of a process by which observations of the 'state' of a system help to constrain the behavior of the system in time. An example of one of the earliest applications of data assimilation is the system in which the trajectory of a flying rocket is constantly (and rapidly) adjusted based on information of its current position to guide it to its exact final destination. Another example of data assimilation is a weather model that gets updated every few hours with measurements of temperature and other variables, to improve the accuracy of its forecast for the next day, and the next, and the

next. Data assimilation is usually a cyclical process, as estimates get refined over time as more observations about the "truth" become available. Mathematically, data assimilation can be done with any number of techniques. For large systems, so-called variational and ensemble techniques have gained most popularity. Because of the size and complexity of the systems studied in most fields, data assimilation projects inevitably include supercomputers that model the known physics of a system. Success in guiding these models in time often depends strongly on the number of observations available to inform on the true system state.

In CarbonTracker, the model that describes the system contains relatively simple descriptions of biospheric and oceanic CO_2 exchange, as well as fossil fuel and fire emissions. In time, we alter the behavior of this model by adjusting a small set of parameters as described in the next section.

2. Detailed Description

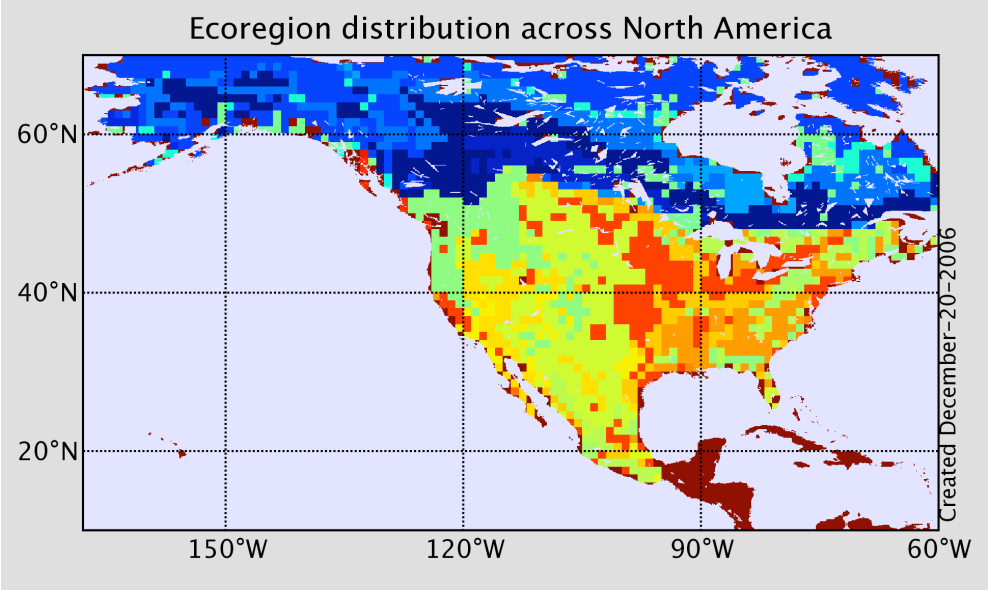
The four surface flux modules drive instantaneous CO_2 fluxes in CarbonTracker according to:

$$F(x, y, t) = \lambda \bullet F_{bio}(x, y, t) + \lambda \bullet F_{oce}(x, y, t) + F_{ff}(x, y, t) + F_{fire}(x, y, t)$$

Where λ represents a set of linear scaling factors applied to the fluxes, to be estimated in the assimilation. These scaling factors are the final product of our assimilation and together with the modules determine the fluxes we present in CarbonTracker. Note that no scaling factors are applied to the fossil fuel and fire modules.

2.1 Land-surface classification

The scaling factors λ are estimated for each week and assumed constant over this period. Each scaling factor is associated with a particular region of the global domain, and currently the geographical distribution of the regions is fixed. The choice of regions is a strong *a-priori* constraint on the resulting fluxes and should be approached with care to avoid so-called "aggregation errors" [Kaminsky, 1999]. We chose an approach in which the ocean is divided up into 11 large basins encompassing large-scale ocean circulation features, as in the TransCom inversion study (e.g. Gurney et al., [2002]). The terrestrial biosphere is divided up according to ecosystem type as well as geographical location. Thereto, each of the 11 TransCom land regions contains a maximum of 19 ecosystem types summarized in the table and figure below. Note that there is currently no requirement for ecoregions to be contiguous, and a single scaling factor can be applied to the same vegetation type on both sides of a continent.



Theoretically, this approach leads to a total number of 11*19+11=220 optimizable scaling factors λ each week, but the actual number is 135 since not every ecosystem type is represented in each TransCom region, and because we decided not to optimize parameters for ice-covered regions, inland water bodies, and deserts. The total flux coming out of these last regions is negligibly small. It is important to note that even though only one parameter is available to scale, for instance, the flux from coniferous forests in Boreal North America, each 1x1 degree grid box predominantly covered by coniferous forests will have a different flux F(x,y,t) depending on local temperature, radiation, and CASA modeled monthly mean flux.

Ecosystem types considered on 1x1 degree for the terrestrial flux inversions is based on Olson, [1992]. Note that we have adjusted the original 29 categories into only 19 regions. This was done mainly to fill the unused categories 16,17, and 18, and to group the from our perspective similar categories 23-26+29. The table below shows each vegetation category considered. Percentages indicate the area associated with each category for North America rounded to one decimal.

Ecosystem Types

category	Olson V 1.3a	Percentage area
1	Conifer Forest	19.0%
2	Broadleaf Forest	1.3%
3	Mixed Forest	7.5%
4	Grass/Shrub	12.6%
5	Tropical Forest	0.3%
6	Scrub/Woods	2.1%

7	Semitundra	19.4%
8	Fields/Woods/Savanna	4.9%
9	Northern Taiga	8.1%
10	Forest/Field	6.3%
11	Wetland	1.7%
12	Deserts	0.1%
13	Shrub/Tree/Suc	0.1%
14	Crops	9.7%
15	Conifer Snowy/Coastal	0.4%
16	Wooded tundra	1.7%
17	Mangrove	0.0%
18	Non-optimized areas (ice, polar desert, inland seas)	0.0%
19	Water	4.9%

Each 1x1 degree pixel of our domain was assigned one of the categories above bases on the Olson category that was most prevalent in the 0.5x0.5 degree underlying area.

2.2 Ensemble Size and Localization

The ensemble system used to solve for the scalar multiplication factors is similar to that in Peters et al. [2005] and based on the square root ensemble Kalman filter of Whitaker and Hamill, [2002]. We have restricted the length of the smoother window to only five weeks as we found the derived flux patterns within North America to be robustly resolved well within that time. We caution the CarbonTracker users that although the North American flux results were found to be robust after five weeks, regions of the world with less dense observational coverage (the tropics, Southern Hemisphere, and parts of Asia) are likely to be poorly observable even after more than a month of transport and therefore less robustly resolved. Although longer assimilation windows, or long prior covariance length-scales, could potentially help to constrain larger scale emission totals from such areas, we focus our analysis here on a region more directly constrained by real atmospheric observations.

Ensemble statistics are created from 150 ensemble members, each with their own background CO_2 concentration field to represent the time history (and thus covariances) of the filter. To dampen spurious noise due to the approximation of the covariance matrix, we apply localization [Houtekamer and Mitchell, 1998] for non-MBL sites only. This ensures that tall-tower observations within North America do not inform on for

instance tropical African fluxes, unless a very robust signal is found. In contrast, MBL sites with a known large footprint and strong capacity to see integrated flux signals are not localized. Localization is based on the linear correlation coefficient between the 150 parameter deviations and 150 observation deviations for each parameter, with a cut-off at a 95% significance in a student's T-test with a two-tailed probability distribution.

2.3 Dynamical Model

In CarbonTracker, the dynamical model is applied to the mean parameter values λ as:

$$\lambda_t^b = (\lambda_{t-2}^a + \lambda_{t-1}^a + \lambda^p) \ / \ 3.0$$

Where "a" refers to analyzed quantities from previous steps, "b" refers to the background values for the new step, and "p" refers to real *a-priori* determined values that are fixed in time and chosen as part of the inversion set-up. Physically, this model describes that parameter values λ for a new time step are chosen as a combination between optimized values from the two previous time steps, and a fixed prior value. This operation is similar to the simple persistence forecast used in Peters et al. [2005], but represents a smoothing over three time steps thus dampening variations in the forecast of λ^b in time. The inclusion of the prior term λ^p acts as a regularization [Baker et al., 2006] and ensures that the parameters in our system will eventually revert back to predetermined prior values when there is no information coming from the observations. Note that our dynamical model equation does not include an error term on the dynamical model, for the simple reason that we don't know the error of this model. This is reflected in the treatment of covariance, which is always set to a prior covariance structure and not forecast with our dynamical model.

2.4 Covariance Structure

Prior values for λ^p are all 1.0 to yield fluxes that are unchanged from their values predicted in our modules. The prior covariance structure P^p describes the magnitude of the uncertainty on each parameter, plus their correlation in space. The latter is applied such that the same ecosystem types in different TransCom regions decrease exponentially with distance (L=2000km), and thus assumes a coupling *between* the behavior of the same ecosystems in close proximity to one another (such as coniferous forests in Boreal and Temperate North America). Furthermore, all ecosystems *within* tropical TransCom regions are coupled decreasing exponentially with distance since we do not believe the current observing network can constrain tropical fluxes on sub-continental scales, and want to prevent large dipoles to occur in the tropics.

In our standard assimilation, the chosen standard deviation is 80% on land parameters, and 40% on ocean parameters. This reflects more prior confidence in the ocean fluxes than in terrestrial fluxes, as is assumed often in inversion studies and partly reflects the lower variability and larger homogeneity of the ocean fluxes. All parameters have the

same variance within the land or ocean domain. Because the parameters multiply the net-flux though, ecosystems with larger weekly mean net fluxes have a larger variance in absolute flux magnitude.

3. Further Reading

- Whitaker and Hamill, 2002 paper
- Peters et al., 2005 paper
- Olson ecosystem types, data

References

- Blasing, T.J., C.T. Broniak, and G. Marland (2005), The annual cycle of fossil-fuel carbon dioxide emissions in the United States, *Tellus B*, 57(2), 107--115
- Giglio, L., d.W. van, G.R., J.T. Randerson, G.J. Collatz, and P. Kasibhatla (2006), Global estimation of burned area using MODIS active fire observations, *Atm. Chem. Phys.*, 6(4), 957-974
- Gurney, K.R. *et al.* (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, 415(6872), 626-630
- Houtekamer, P.L., and H.L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Mon. Weather. Rev.*, 126(3), 796-811
- Jacobson, A.R. *et al.* (2006), A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: II. Results, *submitted to Glob. Biogeochemical Cycles*,
- Kaminski, T., M. Heimann, and R. Giering (1999), A coarse grid three-dimensional global inverse model of the atmospheric transport - 2. Inversion of the transport of CO_2 in the 1980s, *J. Geophys. Res.*, 104(D15), 18555-18581
- Kettle, H., and C.J. Merchant (2005), Systematic errors in global air-sea CO_2 flux caused by temporal averaging of sea-level pressure, *Atm. Chem. Phys.*, 5(6), 1459-1466
- Krol, M.C. *et al.* (2005), The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, *Atm. Chem. Phys.*, 5417-432
- Loveland, T.R. *et al.* (2000), Development of a global land cover characteristics database and IGB6 DISCover from the 1km AVHRR data, *Int. J. Remote Sensing*, 211303-1330
- Olivier, J. G. J., Berdowski, J. J. M. (2001). Global emissions sources and sinks. *In The Climate System*J. Berdowski, R. Guicherit,B. J. Heij, Lisse, The Netherlands: A.A. Balkema Publishers/Swets and Zeitlinger Publishers, pp. 33-78
- Olsen, S.C., and J.T. Randerson (2004), Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, *Journal of Geophysical Research-Atmospheres*, 109(D2),

- Olson, J.S., J.A. Watts, and L.J. Allison (1985), Major world ecosystem complexes ranked by carbon in live vegetation: A Database, *NDP-017*,
- Peters, W. *et al.* (2004), Toward regional-scale modeling using the two-way nested global model TM5: Characterization of transport using SF6, *J. Geophys. Res.*, 109(D19), doi:10.1029/2004JD005020
- Peters, W. *et al.* (2005), An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, *J. Geophys. Res.*, 110(D24304), doi:10.1029/2005JD006157.
- Takahashi, T. *et al.* (2002), Global sea-air CO_2 flux based on climatological surface ocean pCO(2), and seasonal biological and temperature effects, *Deep-Sea Research Part Ii-Topical Studies in Oceanography*, 49(9-10), 1601-1622
- Van Der Werf, G.R. *et al.* (2003), Carbon Emissions from fires in tropical and subtropical ecosystems, *Global Change Biol.*, 9 (4), 547-562
- Van Der Werf, G.R. *et al.* (2004), Continental-Scale Partitioning of Fire Emissions during the 1997 to 2001 El Nino/La Nina Period, *Science*, 303(5654), 73-76
- Van Der Werf, G.R. *et al.* (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atm. Chem. Phys.*, 6(11), 3423-3441
- Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.*, 97(C5), 7373-7382
- Whitaker, J.S., and T.M. Hamill (2002), Ensemble Data Assimilation without Perturbed Observations, *Mon. Weather. Rev.*, 7(130 http://dx.doi.org/10.1175/1520-0493(2002), 1913-1924

Version History

General Release Notes for CarbonTracker 2007

1. CarbonTracker 2007 has a number of caveats for the year 2005 listed below. These shortcomings will most likely lead to revised flux estimates for this year in our next CarbonTracker release (2008). Most importantly, we expect the total uncertainty on the 2005 estimate to become smaller even than the estimate of 2004, as more observations become available for assimilation.
2. The CO_2 measurements obtained from the Meteorological Service Canada (MSC) currently go through February 2005. This record will be expanded for CarbonTracker 2008.
3. The month December of 2005 was run without the United States 1x1 degree zoom region in the transport model. This is due to a change in model resolution in the parent model from ECMWF in February of 2006. We expect future CarbonTracker releases to be able

to integrate past this moment in time as we develop the software to deal with this discontinuity.

4. The 2005 fossil fuel emissions are identical to those used in 2004. Although economic statistics are available to extrapolate the fossil fuel emissions from previous years, we did not include this information in version 2007. We expect better fossil fuel emission estimates for 2005 to be included in CarbonTracker 2008
5. The 2005 fire emissions are the climatological average of those used in 2000-2004. The new GFED2 release (weekly fluxes through 2005) came out too late to be incorporated in CarbonTracker 2007
6. Annual mean uncertainties for CarbonTracker 2007 are calculated as the annual average of the 52 weekly estimated covariances. This estimate does not include any temporal covariances in the fluxes which would substantially reduce uncertainty. We feel that until we have implemented a proper mechanism to propagate such information through time, we can not objectively quote the true uncertainty on the annual mean that would result from the assimilation of all the observations. Our current number is supposed to represent a conservative estimate of the real uncertainty to be compared to previous flux estimates.
7. Uncertainty maps on the fluxes show smaller uncertainties in the tropics than in the extra-tropics. This is counter-intuitive as the tropical carbon cycle is least well understood and most poorly constrained by observations. In recognition of this, CarbonTracker 2007 uses strong covariances on ecosystems in the tropical regions. The variance on each region is lowered to limit the resulting total uncertainty per continent. Since our maps only show variances per region, and not covariances, this fact is not reflected in the mapped product.
8. Fossil fuel emissions can occur over regions characterized as ocean. This is partly due to real emissions from international shipping, and partly due to emissions occurring in coastal land regions that are assigned to the ocean in our coarse 1x1 degree aggregation scheme. The same is true for fossil fuel emissions over non-optimized regions such as ice, polar deserts, and inland seas. This problem shows up only when aggregating to larger land, ocean, or continental regions for interpretation, but has no influence on the assimilation process itself.