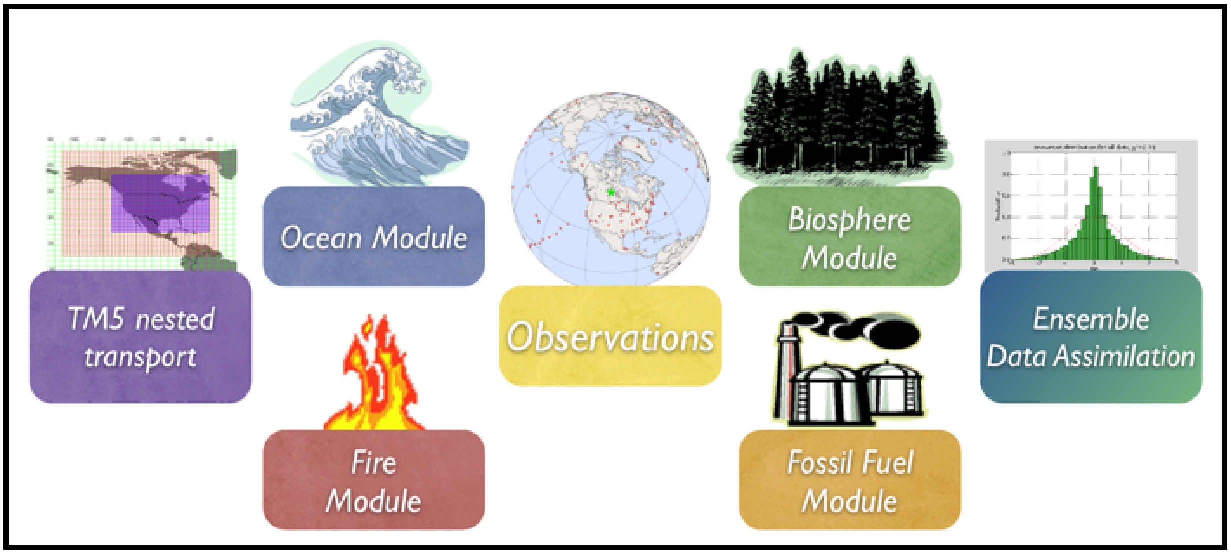
Documentation - CT2011_oi



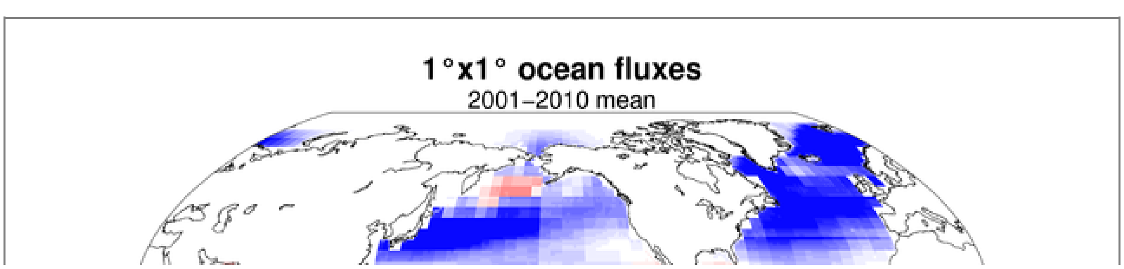
To learn more about a CarbonTracker component, click on one of the above images.
Or download the full PDF version for convenience.

Oceans Module [goto top]

Notice: CT2011_oi does not use the climatological ocean prior discussed here. After we released CarbonTracker 2011, a significant bug was discovered in our atmospheric transport model. We have corrected the bug and are releasing revised results under the release name "CT2011_oi". One consequence of this problem is that the four inversions using the climatological ocean flux prior were faulty. They have been removed from the inversion suite in CT2011_oi. Use of the original CT2011 results is strongly discouraged. Details can be found at this link. While it is not used in CT2011_oi, documentation regarding the climatological ocean prior is retained here as a reference.

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 Earth system. Oceanographers estimate that about 48% of the CO_2 from fossil fuel burning has been absorbed by the ocean [Sabine et al., 2004]. The dissolution of CO_2 in seawater shifts the balance of the ocean carbonate equilibrium towards a more acidic state with a lower pH. This effect is already measurable [Caldeira and Wickett, 2003], and is expected to become an acute challenge to shell-forming organisms over the coming decades and centuries. Although the oceans as a whole have been a relatively steady net carbon sink, CO_2 can also be released from oceans depending on local temperatures, biological activity, wind speeds, and ocean circulation. These processes are all considered in CarbonTracker, since they can have significant effects on the ocean sink. Improved estimates of the air-sea exchange of carbon in turn help us to understand variability of both the atmospheric burden of CO_2 and terrestrial carbon exchange.



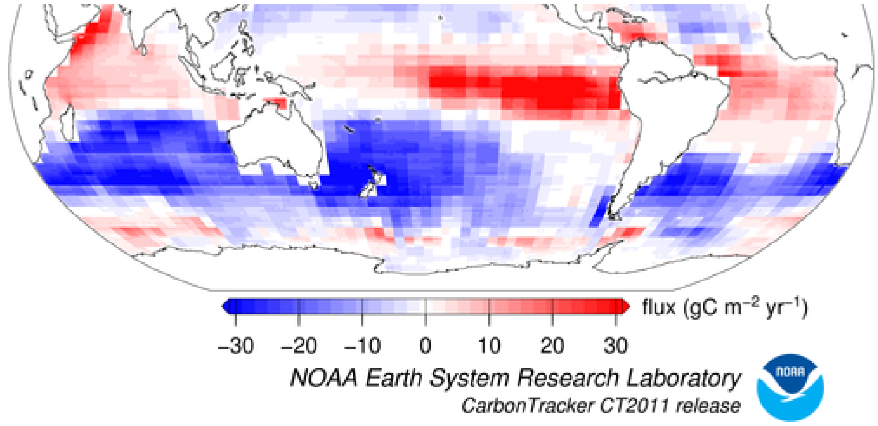


Figure 1. Posterior long-term mean ocean fluxes from CarbonTracker. The pattern of air-sea exchange of CO_2 averaged over the time period indicated, as estimated by CarbonTracker. Negative fluxes (blue colors) represent CO_2 uptake by the ocean, whereas positive fluxes (red colors) indicate regions in which the ocean is a net source of CO_2 to the atmosphere. Units are gC m^{-2} yr^{-1}.

The initial release of CarbonTracker (2007) used climatogical estimates of CO_2 partial pressure in surface waters (pCO_2) from Takahashi et al. [2002] to compute a first-guess air-sea flux. This air-sea pCO_2 disequilibrium was modulated by a surface barometric pressure correction before being multiplied by a gas-transfer coefficient to yield a flux. Starting with CarbonTracker 2007B and continuing through the CT2010 release, the air-sea pCO_2 disequilibrium was imposed from analysis of ocean inversions ("OIF", cf. Jacobson et al., 2007) results, with short-term flux variability derived from the atmospheric model wind speeds via the gas transfer coefficient. The barometric pressure correction was removed so that climatological high- and low-pressure cells did not bias the long-term means of the first guess fluxes.

Starting with the CT2011 release, two models are used to provide prior estimates of air-sea CO_2 flux. The OIF scheme provides one of these flux priors, and the other is an updated version of the Takahashi et al. pCO_2 climatology.

2. Detailed Description

Oceanic uptake of CO_2 in CarbonTracker is computed using air-sea differences in partial pressure of CO_2 inferred either from ocean inversions (called "OIF" henceforth), or from a compilation of direct measurements of seawater pCO_2 (called "pCO_2-clim" henceforth). These air-sea partial pressure differences are combined with a gas transfer velocity computed from wind speeds in the atmospheric transport model to compute fluxes of carbon dioxide across the sea surface.

In either method, the first-guess fluxes have no interannual variability (IAV) other than a smooth trend. IAV in oceanic CO_2 flux is due to anomalies in surface pCO_2, such as those that occur in the tropical eastern Pacific during an El Niño, and to associated variability in winds, ocean circulation, and sea-surface properties. In CarbonTracker, only the surface winds (and hence gas transfer), manifest these interannual anomalies; the remaining IAV of flux must be inferred from atmospheric CO_2 signals.

In the following sections we describe the two ocean flux prior models. We then describe the air-sea gas transfer velocity parameterization and discuss detais of the inversion methodology specific to oceanic exchange of CO_2. Figures comparing the air-sea flux priors are presented in Box 1 below.

2.1 OIF: the Ocean Inversion Fluxes prior

For the OIF prior, long-term mean air-sea fluxes and the uncertainties associated with them are derived from the ocean interior inversions reported in Jacobson et al. [2007]. These ocean inversion flux estimates are composed of separate preindustrial (natural) and anthropogenic flux inversions based on the methods described in Gloor et al. [2003] and biogeochemical interpretations of Gruber, Sarmiento, and Stocker [1996]. The uptake of anthropogenic CO_2 by the ocean is assumed to increase in proportion to atmospheric CO_2 levels, consistent with estimates from ocean carbon models.

OIF contemporary pCO_2 fields were computed by summing the preindustrial and anthropogenic flux components from inversions using five different configurations of the Princeton/GFDL MOM3 ocean general circulation model [Pacanowski and Gnanadesikan, 1998], then dividing by a gas transfer velocity computed from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA40 reanalysis. There are two small differences in first-guess fluxes in this computation from those reported in Jacobson et al. [2007]. First, the five OIF estimates all used Takahashi et al. [2002] pCO_2 estimates to provide high-resolution patterning of flux within inversion regions (the alternative "forward" model patterns were not used). To good approximation, this choice only affects the spatial and temporal distribution of flux within each of the 30 ocean inversion regions, not the magnitude of the estimated flux. Second, wind speed differences between the ERA40 product used in the offline analysis and the ECMWF operational model used in the online CarbonTracker analysis result in small deviations from the OIF estimates.

Other than the smooth trend in anthropogenic flux assumed by the OIF results, interannual variability (IAV) in the first guess ocean flux comes entirely from wind speed effects on the gas transfer velocity. This is because the ocean inversions retrieve only a long-term mean and smooth trend.

2.2 pCO_2-Clim: Takahashi et al. [2009] climatology prior

The pCO_2-Clim prior is derived from the Takahashi et al. [2009] climatology of seawater pCO_2. This climatology was created from about 3 million direct observations of seawater pCO_2 around the world between 1970 and 2007. With the exception of measurements in the Bering Sea, these observations were all linearly extrapolated to the corresponding month of the year 2000 by assuming a constant trend of 1.5 $\mu atm\ yr^{-1}$. This set of global monthly measurements corrected to the reference year 2000 was then interpolated onto a regular grid using a modeled surface current field.

The Takahashi et al. [2009] product goes beyond providing this estimate of surface water pCO_2. They also compute climatological air-sea exchange of CO_2 by using the GLOBALVIEW-CO_2 atmospheric carbon dioxide product to infer air-sea ΔpCO_2, sea surface properties inferred from ocean climatologies, and winds from atmospheric reanalysis to estimate gas-transfer velocity. Unlike other atmospheric analyses, we have chosen not to use the climatological fluxes as our prior, nor to use the climatological ΔpCO_2.

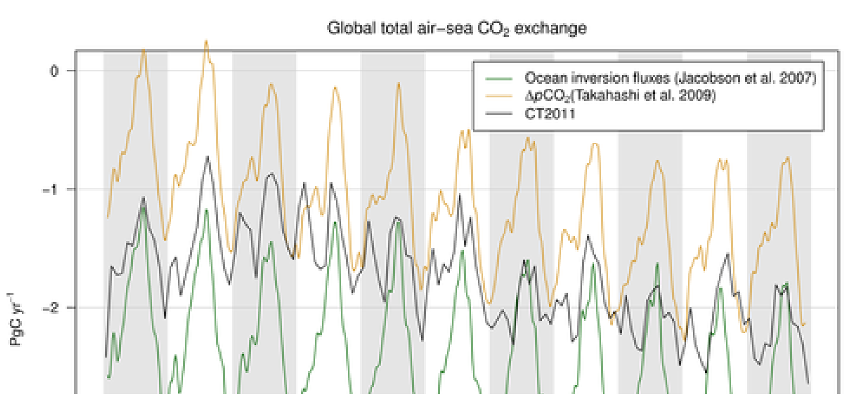
Instead, we take only the seawater pCO_2 distribution from the Takahashi et al. climatology--our atmospheric model provides both pCO_2 in the air at the sea surface and the winds needed to estimate gas transfer. Seawater pCO_2 is extrapolated from 2000 to the actual year of the CarbonTracker simulation using the presumed increase of 1.5 $\mu atm\ yr^{-1}$ at every point in the global ocean.

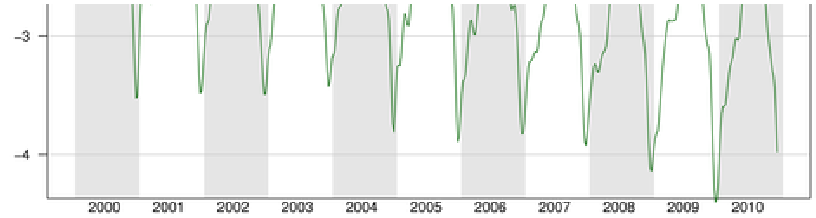
2.3 Gas-transfer velocity and ocean surface properties

Both priors use CO_2 solubilities and Schmidt numbers computed from World Ocean Atlas 2009 (WOA09) climatological fields of sea surface temperature [Locarnini et al., 2010] and sea surface salinity [Antonov et al., 2010] fields. Gas transfer velocity in CarbonTracker is parameterized as a quadratic function of wind speed following Wanninkhof [1992], using the formulation for instantaneous winds. Gas exchange is computed every 3 hours using wind speeds from the ECMWF operational model as represented by the TM5 atmospheric transport model.

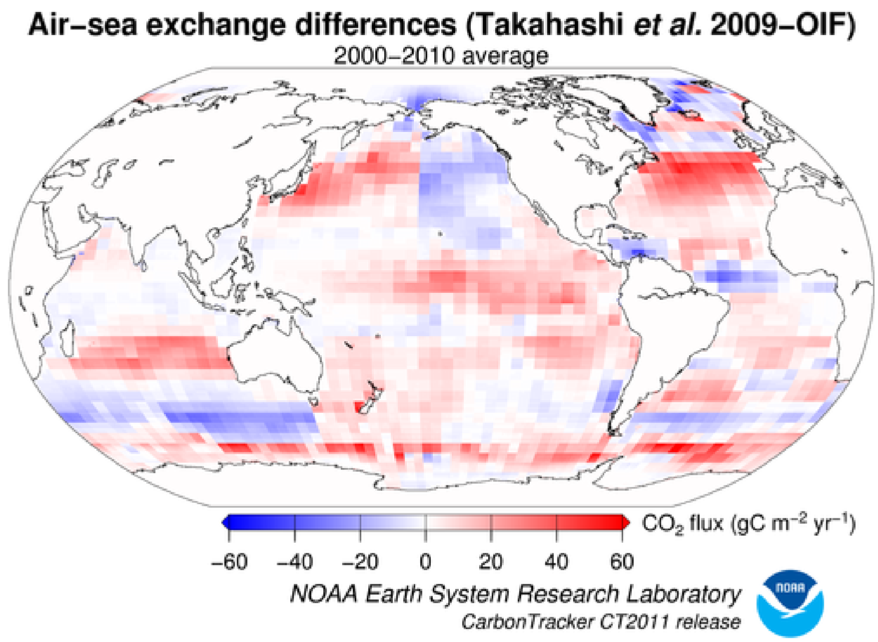
Air-sea transfer is inhibited by the presence of sea ice, and for this work fluxes are scaled by the daily sea ice fraction in each gridbox provided by the ECMWF forecast data.

Box 1. Comparison of air-sea flux priors





Time series of global-total ocean flux among the two priors and the CT2011 posterior. Global CO_2 uptake by the ocean, expressed in PgC yr^{-1}. Positive flux represents a gain of CO_2 to the atmosphere, and the negative numbers here indicate that the ocean is a sink of CO_2. While both priors manifest similar trends of increasing oceanic uptake of CO_2, the OIF prior (in green) has more oceanic uptake and a greater annual cycle than the pCO_2-clim prior (in tan). The CT2011 across-model posterior estimate is shown in black for comparison.



Differences in long-term mean ocean fluxes between the two priors. Red indicates areas where the pCO_2-clim prior has less oceanic uptake (or more outgassing to the atmosphere) than the OIF prior, and blue represents the opposite. Units are $gC\ m^{-2}\ yr^{-1}$.

2.4. Specifics of the Inversion Methodology Specific to Air-sea CO_2 Fluxes

The first-guess fluxes described here are subject to scaling during the CarbonTracker optimization process, in which atmospheric CO_2 mole fraction observations are combined with transport simulated by the atmospheric model to infer flux signals. Prior air-sea fluxes are adjusted within each of of the 30 ocean inversion regions. In this process, signals of terrestrial flux in atmospheric CO_2 distribution can be erroneously interpreted as being caused by oceanic fluxes. This flux "aliasing" or "leakage" is evident in some regions as a change in the shape of the seasonal cycle of air-sea flux.

Prior uncertainties for the OIF and pCO_2-clim models are specified as uncertainties on scaling factors multiplying net CO_2 flux in each of the 30 ocean inversion regions. The pCO_2-clim prior has independent regional uncertainties (a diagonal prior covariance matrix), with the uncertainty standard deviation on each region set to 40%. The OIF prior uncertainty has a fully-covariate covariance matrix with off-diagonal elements representing the results of the ocean inversion of Jacobson et al. [2007]. The pre-industrial flux uncertainty is time-indepedendent, but the anthropogenic flux uncertainty grows in time as anthropogenic flux uptake increases. The latter is scaled to the simulation date, then added to the former. Total uncertainties are consistent with the Jacobson et al. [2007] results.

3. Further Reading

- NOAA Pacific Marine Environmental Laboratory (PMEL)

- Ocean Acidification
- Antonov, J. I., D. Seidov, T. P. Boyer, R. A. Locarnini, A. V. Mishonov, H. E. Garcia, O. K. Baranova, M. M. Zweng, and D. R. Johnson, 2010. World Ocean Atlas 2009, Volume 2: Salinity. S. Levitus, Ed. NOAA Atlas NESDIS 69, U.S. Government Printing Office, Washington, D.C., 184 pp.
- Caldeira, K., and M. E. Wickett (2003), Anthropogenic carbon and ocean pH, Nature, 425365-365.
- GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA ESRL, Boulder, Colorado [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW], 2011.
- Gloor, M., N. Gruber, J. Sarmiento, C. L. Sabine, R. A. Feely, and C. Rödenbeck (2003), A first estimate of present and preindustrial air-sea CO_2 flux patterns based on ocean interior carbon measurements and models, Geophysical Research Letters, 30, 10.1029/2002GL015594.
- Gruber, N., J. L. Sarmiento, and T. F. Stocker (1996), An improved method for detecting anthropogenic CO_2 in the oceans, Global Biogeochemical Cycles, 10, , 809-837.
- Jacobson, A. R., N. Gruber, J. L. Sarmiento, M. Gloor, and S. E. Mikaloff Fletcher (2007), A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: I. Methods and global-scale fluxes, Global Biogeochemical Cycles, 21, doi:10.1029/2005GB002556.
- Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, and D. R. Johnson, 2010. World Ocean Atlas 2009, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 68, U.S. Government Printing Office, Washington, D.C., 184 pp.
- Pacanowski, R. C., and A. Gnanadesikan (1998), Transient response in a z-level ocean model that resolves topography with partial cells, Monthly Weather Review, 126, 3248--3270.
- Sabine, C. L., R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tilbrook, F. J. Millero, T. H. Peng, A. Kozyr, T. Ono, and A. F. Rios (2004), The oceanic sink for anthropogenic CO_2, Science, 305, 367-371.
- Takahashi, T., S. C. Sutherland, C. Sweeney, A. P. N. Metzl, B. Tilbrook, N. Bates, R. Wanninkhof, R. A. Feely, C. Sabine, J. Olafsson, and Y. Nojiri (2002), Global air-sea CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, Deep-Sea Research II, 49, 1601--1622.
- Takahashi, T., S. C. Sutherland, C. Sweeney, R. A. Feely, D. W. Chipman, B. Hales, G. Friederich, F. Chavez, C. Sabine, A. Watson, D. C. E. Bakker, U. Schuster, N. Metzl, H. Yoshikawa-Inoue, M. Ishii, T. Midorikawa, Y. Nojiri, A. Kortzinger, T. Steinhoff, M. Hoppema, J. Olaffson, T. S. Anarson, B. Tilbrook, T. Johannessen, A. Olsen, R. Bellerby, C. S. Wong, B. Delille, N. R. Bates, and H. J. W. de Baar (2009), Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, Deep-Sea Resarch II, 56, 554--577.
- Wanninkhof, R. (1992), Relationship between wind speed and gas exchange over the ocean, Journal of Geophysical Research, 97, 7373--7382.

Terrestrial Biosphere Module [goto top]

1. Introduction

The biospheric component of the terrrestrial 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 our fire emissions module documentation. 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 PgC 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 0.5° x 0.5° 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.

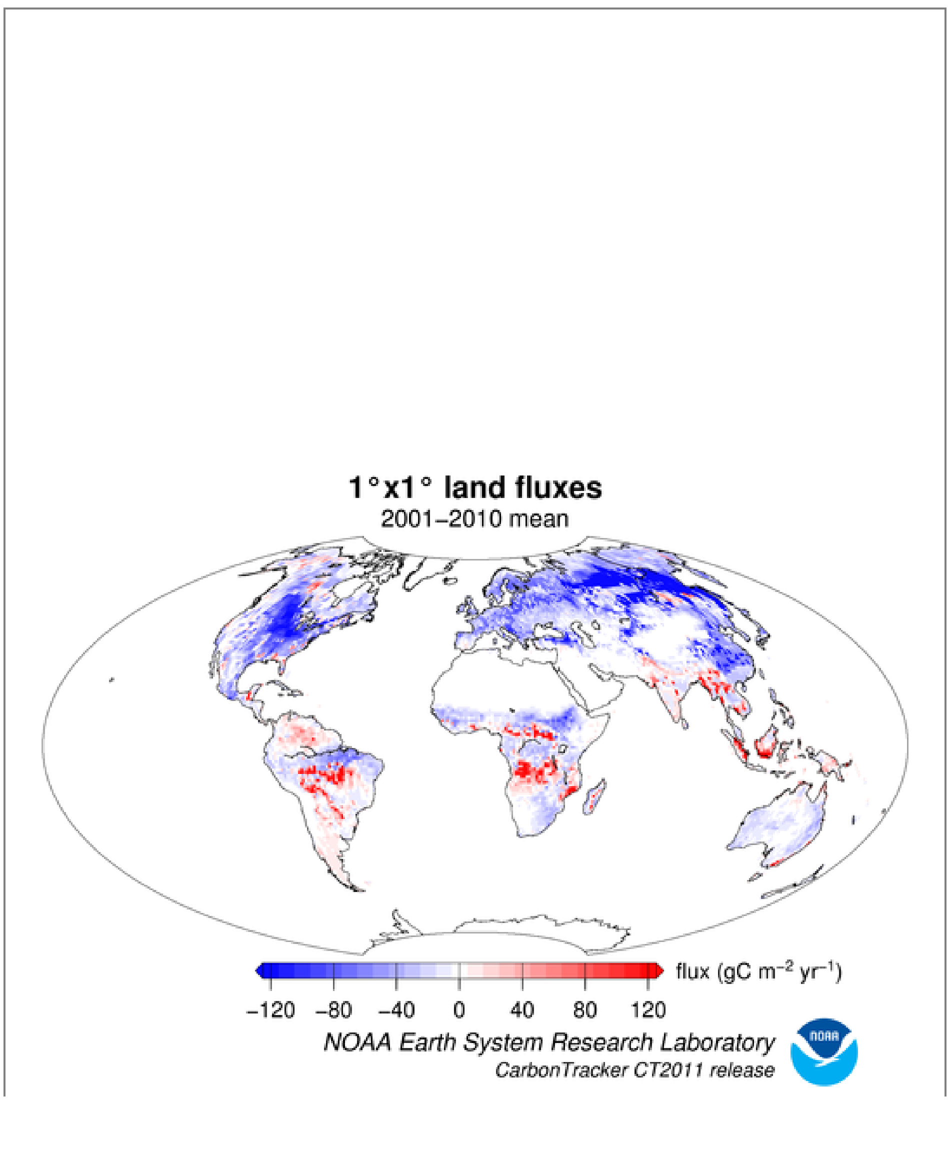
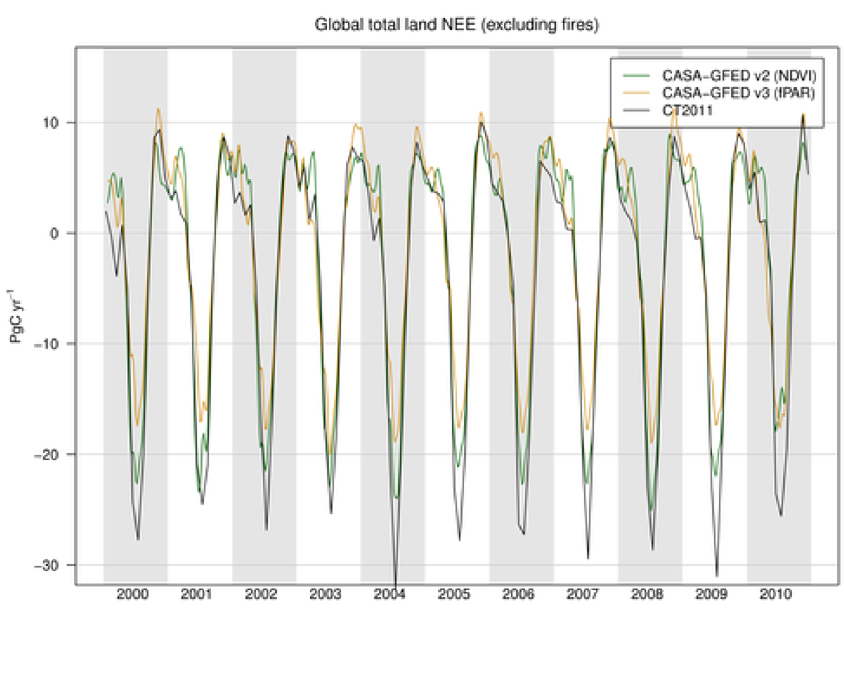


Fig 1. Map of optimized global biosphere fluxes. The pattern of net ecosystem exchange (NEE) of CO_2 of the land biosphere averaged over the time period indicated, as estimated by CarbonTracker. This NEE represents land-to-atmosphere carbon exchange from photosynthesis and respiration in terrestrial ecosystems, and a contribution from fires. It does not include fossil fuel emissions. Negative fluxes (blue colors) represent CO_2 uptake by the land biosphere, whereas positive fluxes (red colors) indicate regions in which the land biosphere is a net source of CO_2 to the atmosphere. Units are gC m^{-2} yr^{-1}.

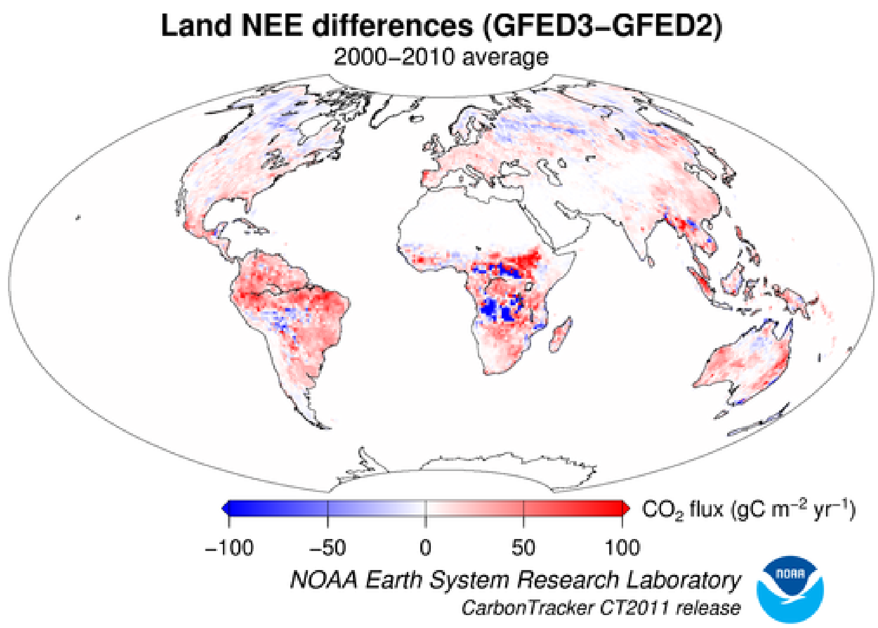
CarbonTracker uses fluxes from CASA runs for the GFED project as its first guess for terrestrial biosphere fluxes. We have found a significantly better match to observations when using this output compared to the fluxes from a neutral biosphere simulation. Prior to CT2010, we used version 2 of the CASA-GFED model, which is driven by AVHRR NDVI, scaled to represent MODIS fPAR. Recently the GFED team has transitioned to version 3.1 of their model, driven directly by MODIS fPAR. We have found that the newer CASA-GFEDv3 product has a smaller seasonal cycle than the older CASA-GFEDv2.

The record of atmospheric CO_2 calls for a deeper terrestrial biosphere sink than that generally simulated by forward models like CASA-GFED. This is manifested by a larger annual cycle of terrestrial biosphere fluxes, and in particular a deeper boreal summer uptake of carbon dioxide, in the posterior optimized fluxes compared to the prior models (See Box 1). We call upon the atmospheric CO_2 observations to make this change, and in order to handle these prior model differences the ensemble Kalman filter's prior covariance model has been re-tuned. In short, this prior uncertainty needs to comfortably span differences among the terrestrial biosphere priors, the fossil fuel emissions priors, and adjustments to fluxes required to bring model predictions into agreement with observations. As a result, the land biosphere prior uncertainty has been doubled in CT2011 in comparison to previous releases. Details can be found on the assimilation scheme documentation page.

Box 1. Comparison of terrestrial biosphere flux priors



Time series of global-total terrestrial biosphere flux between the two priors and the CT2011 posterior. Global CO_2 uptake by the land biosphere, expressed in PgC yr^{-1}, excluding emissions by wildfire. Positive flux represents emission of CO_2 to the atmosphere, and the negative fluxes indicate times when the land biosphere is a sink of CO_2. While both priors manifest similar annual cycles of uptake in boreal summer balanced by emission in boreal winter, the GFED3 prior (tan) has an annual cycle that is about 10% smaller than that of GFED2 (green). Optimization against atmospheric CO_2 data requires a larger land sink than in either prior, which effectively requires a deeper annual cycle. This is shown by the CT2011 posterior (black).



Differences in long-term mean terrestrial biosphere fluxes between the two priors. Red indicates areas where the GFED3 prior has less terrestrial uptake (or more outgassing to the atmosphere) than the GFED2 prior, and blue represents the opposite. Units are gC m^{-2} yr^{-1}.

Unlike CT2010, CarbonTracker 2011 is a full reanalysis of the 2000-2010 period using new fossil fuel emissions, CASA-GFEDv3 fire emissions, and first-guess biosphere model fluxes derived from CASA-GFEDv2 for 4 of our inversions, and from CASA-GFEDv3 for the remaining 4 inversions.

Due to the inclusion of fires, inter-annual variability in weather and NDVI (or fPAR), 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^{-1} of release, to 0.15 PgC yr^{-1} 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

Fire Module [goto top]

1. Introduction

Vegetation 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, vegetation fires emit around 2 PgC as CO_2 into the atmosphere, mostly in the tropics. Currently, 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 amount of CO_2 emitted by each fire. 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 Fire Emissions Database (GFED), which uses the CASA biogeochemical model as described in the terrestrial biosphere model documentation to estimate the carbon fuel in various biomass pools. The dataset consists of 1° x 1° 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, Total Particulate Matter, Total Carbon, Organic Carbon, Black Carbon) for the time period spanning January 1997 - December 2009, of which we currently only use CO_2.

In 2010, the GFED team switched the satellite product driving the CASA terrestrial productivity submodel from AVHRR NDVI to the MODIS fPAR product. For CT2011, we use fire emissions from the fPAR-driven GFED 3.1 for the entire simulation period of 2000-2010.

The GFED 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 MODIS in the period 2001-2004.

Once burned area has been estimated globally, 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.

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
- Interannual variability in global biomass burning emissions from 1997 to 2004, G. R. van der Werf, J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr., Atmospheric Chemistry and Physics 6: 3423-3441 Aug 21 2006.

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, amount and location 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.

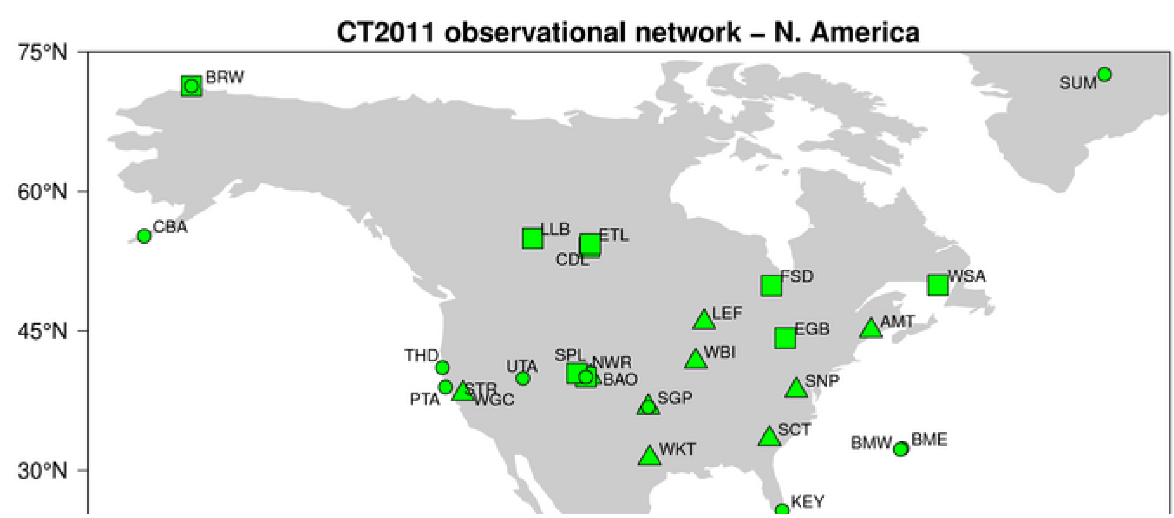
2. Detailed Description

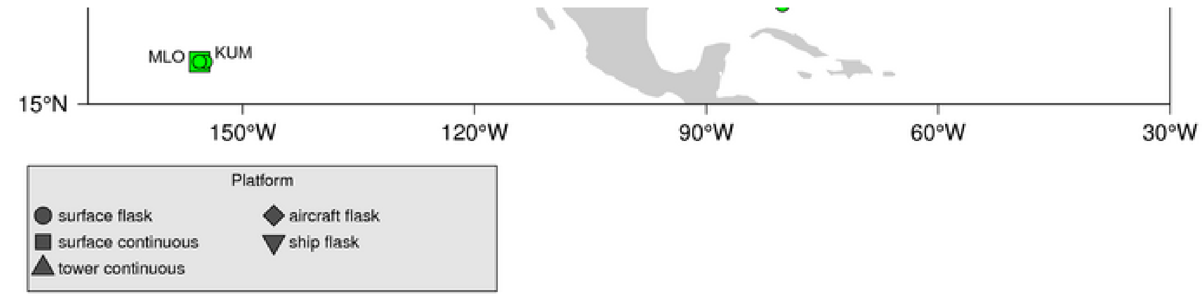
This study uses measurements of air samples collected at surface sites in the NOAA ESRL Cooperative Global Air Sampling Network, the CSIRO Air Sampling Network and the IPEN-CQMA sampling program where available, except those flagged for analysis or sampling problems, or those thought to be influenced by local sources. The sites for which data are available thus varies each week depending on successful sampling and analysis, and each site's sampling frequency. In addition, we use in situ quasi-continuous CO_2 time series from the following towers:

- the 107m level of the AMT tower in Argyle, Maine
- the 300m level of the BAO tower in Boulder, Colorado
- the 396m level of the LEF tower in Park Falls, Wisconsin
- the 305m level of the SCT tower in Beech Island, South Carolina
- the 17m level of the SNP tower in Shenandoah National Park, Virginia
- the 379m level of the WBI tower in West Branch, Iowa
- the 483m level of the WGC tower in Walnut Grove, California
- the 457m level of the WKT tower in Moody, Texas

- the 30m level of the tower at Candle Lake (CDL, formerly Old Black Spruce), Saskatchewan, Canada operated by Environment Canada (EC);
- the 105m level of the tower in East Trout Lake, Saskatchewan, Canada (ETL) operated by EC
- the 40m level of the tower in Fraserdale, Ontario, Canada (FRD) operated by EC
- the 10m level of the tower in Lac Labiche, Alberta, Canada (LLB) operated by EC

- the 60m level of the tower at the Atmospheric Radiation and Monitoring (ARM) Carbon Project Southern Great Plains, Oklahoma site (SGP) operated by Lawrence Berkeley National Laboratory (LBNL).





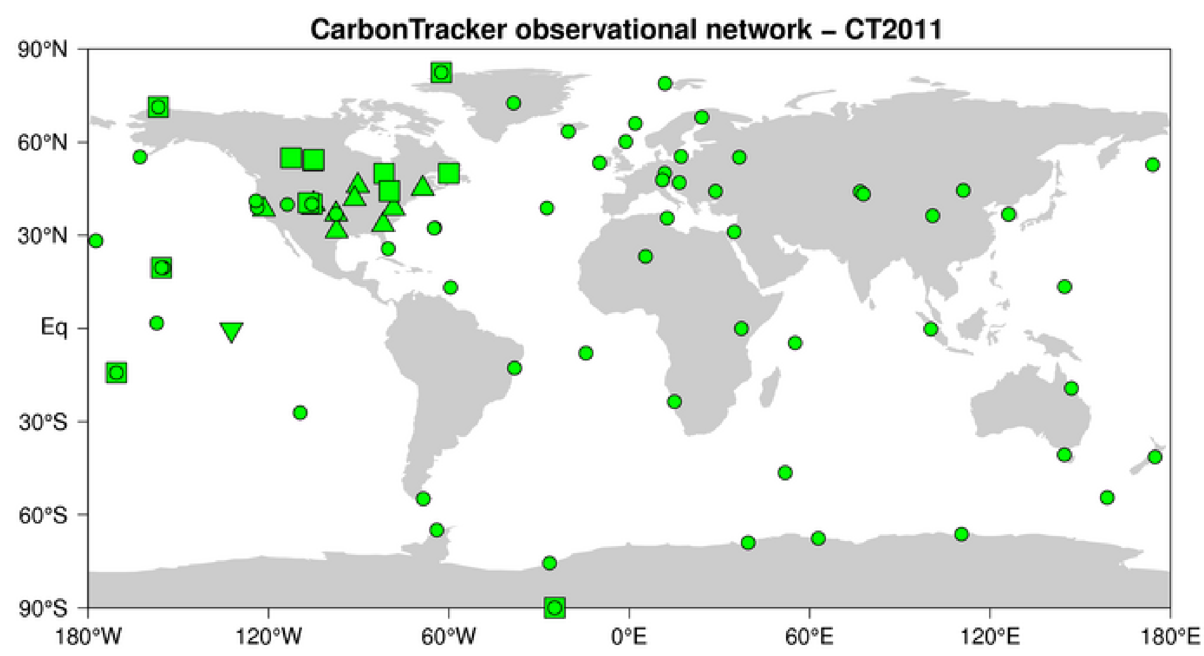
Other in situ quasi-continuous CO_2 time series used are from the NOAA ESRL observatories at Barrow (BRW), Mauna Loa (MLO), Samoa (SMO), and South Pole (SPO); the EC Canadian sites at Alert, Nunavut (ALT), Sable Island, Nova Scotia (SBL) and Egbert, Ontario (EGB); and the NCAR sites at Niwot Ridge, Colorado (NWR) and Storm Peak Laboratory, Colorado (SPL). Note that all of these observations are calibrated against the same world CO_2 standard (WMO-X2007). Also, note that aircraft observations from the NOAA ESRL program were NOT assimilated, but used for independent assessment of the CarbonTracker results.

For most of the quasi-continuous sampling sites, we construct an afternoon daytime average mole fraction for each day from the time series, recognizing that our atmospheric transport model does not always capture the continental nighttime stability regime while daytime well-mixed conditions are better matched. At mountain-top sites (MLO, NWR, and SPL), we use an average of nighttime hours as this tends to be the most stable time period and avoids periods of upslope flows that contain local vegetative and/or anthropogenic influence. Moreover, observations at sub-daily time scales are likely to be strongly correlated and therefore add relatively little independent information to our results.

Data from the Sutro tower (STR_01P0) and the Boulder tower (BAO_01P0, BAO_01C3) are strongly influenced by local urban emissions, which CarbonTracker is unable to resolve. At these two sites, pollution events have been identified using co-located measurements of carbon monoxide. In this study, measurements thought to be affected by pollution events have been excluded. This technique is still being developed.

Also based on Transcom 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. Table 1 summarizes how data from the different measurement programs are preprocessed for this study.

The preprocessed data used in CarbonTracker are freely available for download. Preprocessed data are **not** the original measurement data. Users are encouraged to review the literature and contact the measurement labs directly for details about and access to the actual observations.



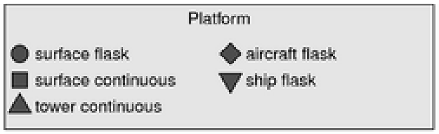


Table 1: Summary of CarbonTracker data preprocessing.

Measurement Program	Data Preprocessing
ESRL discrete surface	All valid[1] data. Multiple values from the same day and location are averaged. No sample time-of-day restriction (see exception below).
ESRL discrete tower	All valid[1] data. Multiple values from the same day and location are averaged. Only samples collected between 12-16 LST are considered.
ESRL observatories (BRW, SMO, SPO)	All valid[1] data. Day average using 12-16 LST.
ESRL observatories (MLO)	All valid[1] data. Day average using 0-4 LST.
ESRL tower sites	All valid data from highest intake. Day average using 12-16 LST.
EC in situ sites	All valid data from highest intake. Day average using 12-16 LST.
NCAR in situ sites	All valid data from highest intake where 1σ of hourly average < 1 ppm. Day average using 0-4 LST.
CSIRO discrete surface	All valid[1] data. Multiple values from the same day and location are averaged. No sample time-of-day restriction.
IPEN discrete surface	All valid[1] data. Multiple values from the same day and location are averaged. No sample time-of-day restriction.
LBNL in situ site	All valid data for the period 2003-2010. Day average using 14-18 LST.

[1]In this context "Valid Data" means the observation is thought to be free of sampling and analytical problems and has not been locally influenced.

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 ascribed 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 1° x 1° 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 2 (below) gives a summary of the observing sites used in CarbonTracker, and the performace of the assimilation scheme at each site. These diagnostics are useful for evaluating how well CarbonTracker does in simulating observed CO_2.

Table 2. Summary of Observational Sites Used in CarbonTracker. Model-data-mismatch ("r") is a value assigned to a given site that is meant to quantify our expected ability to simulate observations there. This value is principally determined from the limitations of the atmospheric transport model. It is part of the standard deviation used to interpret the difference between a simulation first guess ("Hx") of an observation and the actual measured value ("z"). The other component, HPH^T, is a measure of the ability of the ensemble Kalman filter to improve its simulated value for this observation by adjusting fluxes. These elements together form the innovation χ statistic for the site: $\chi = (z-Hx)/\sqrt{(HPH^T+r^2)}$. The innovation χ^2 reported above is the mean of all squared χ values for a given site. An average χ^2 below 1.0 indicates that the $\sqrt{(HPH^T+r^2)}$ values are too large. Conversely, values above 1.0 mean that this standard deviation is underestimated. The bias and SE columns are statistics of the posterior residuals (final modeled values - measured values). The bias is the mean of these residuals; the SE is the standard error of those residuals.

Site code	Lab.	Location	Latitude	Longitude	Elev. (m ASL)	No. Obs. Avail.	No. Obs. Used	No. Obs. Rej.	r ($\mu mol\ mol^{-1}$)	Innov. χ^2	Bias ($\mu mol\ mol^{-1}$)	SE ($\mu mol\ mol^{-1}$)
ABP_01D0	ESRL	Arembepe, Bahia, Brazil	12.77°S	38.17°W	1	102	102	0	2.50	0.06	-0.50	0.53
ABP_26D0	IPEN	Arembepe, Bahia, Brazil	12.77°S	38.17°W	1	101	101	0	2.50	0.11	-0.63	0.68
ALT_01D0	ESRL	Alert, Nunavut,	82.45°N	62.51°W	200	530	530	0	1.50	0.53	0.29	0.90

		Canada										
ALT_06C0_14LST	EC	Alert, Nunavut, Canada	82.45°N	62.51°W	200	3652	3652	0	2.50	0.23	0.29	0.99
AMT_01C3_14LST	ESRL	Argyle, Maine, United States	45.03°N	68.68°W	50	2272	2218	54	3.00	0.93	0.55	3.22
AMT_01P0	ESRL	Argyle, Maine, United States	45.03°N	68.68°W	50	325	310	15	3.00	1.23	0.85	3.52
ASC_01D0	ESRL	Ascension Island, United Kingdom	7.97°S	14.40°W	74	915	915	0	0.75	0.82	−0.01	0.58
ASK_01D0	ESRL	Assekrem, Algeria	23.18°N	5.42°E	2728	480	480	0	1.50	0.34	0.13	0.92
AZR_01D0	ESRL	Terceira Island, Azores, Portugal	38.77°N	27.38°W	40	333	329	4	1.50	1.11	0.62	1.46
BAL_01D0	ESRL	Baltic Sea, Poland	55.35°N	17.22°E	3	929	929	0	7.50	0.26	−1.09	3.90
BAO_01C3_14LST	ESRL	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1584	964	952	12	3.00	1.00	−0.98	2.70
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (µmol mol^{-1})**	**Innov. χ^2**	**Bias (µmol mol^{-1})**	**SE (µmol mol^{-1})**
BAO_01P0	ESRL	Boulder Atmospheric Observatory, Colorado, United States	40.05°N	105.00°W	1584	720	694	26	3.00	1.36	−0.97	3.42
BHD_01D0	ESRL	Baring Head Station, New Zealand	41.41°S	174.87°E	85	142	142	0	1.50	0.28	0.20	0.74
BKT_01D0	ESRL	Bukit Kototabang, Indonesia	0.20°S	100.32°E	864	247	247	0	7.50	0.63	4.55	3.62
BME_01D0	ESRL	St. Davids Head, Bermuda, United Kingdom	32.37°N	64.65°W	30	236	225	11	1.50	1.57	0.69	1.83
BMW_01D0	ESRL	Tudor Hill, Bermuda, United Kingdom	32.27°N	64.88°W	30	347	344	3	1.50	0.87	0.59	1.28
BRW_01C0_14LST	ESRL	Barrow, Alaska, United States	71.32°N	156.61°W	11	2975	2973	2	2.50	0.23	0.07	1.04
BRW_01D0	ESRL	Barrow, Alaska, United States	71.32°N	156.61°W	11	499	497	2	1.50	0.55	−0.01	1.05
BSC_01D0	ESRL	Black Sea, Constanta, Romania	44.17°N	28.68°E	3	399	386	13	7.50	1.18	−5.28	7.32
CBA_01D0	ESRL	Cold Bay, Alaska, United States	55.21°N	162.72°W	21	736	702	34	1.50	1.59	−0.15	1.83
CDL_06C0_14LST	EC	Candle Lake, Saskatchewan, Canada	53.99°N	105.12°W	600	2509	2494	15	3.00	0.59	0.24	2.15
CFA_02D0	CSIRO	Cape Ferguson, Queensland, Australia	19.28°S	147.06°E	2	205	205	0	2.50	0.24	−0.44	1.09
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (µmol mol^{-1})**	**Innov. χ^2**	**Bias (µmol mol^{-1})**	**SE (µmol mol^{-1})**
CGO_01D0	ESRL	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94	404	404	0	0.75	0.16	0.07	0.28
		Cape Grim										

CGO_02D0	CSIRO	Cape Grim, Tasmania, Australia	40.68°S	144.69°E	94	404	404	0	0.75	0.16	0.03	0.27
CHR_01D0	ESRL	Christmas Island, Republic of Kiribati	1.70°N	157.17°W	3	419	419	0	0.75	0.65	−0.38	0.47
CRZ_01D0	ESRL	Crozet Island, France	46.45°S	51.85°E	120	402	402	0	0.75	0.19	0.01	0.27
CYA_02D0	CSIRO	Casey, Antarctica, Australia	66.28°S	110.52°E	51	200	200	0	0.75	0.30	−0.21	0.27
EGB_06C0_14LST	EC	Egbert, Ontario, Canada	44.23°N	79.78°W	251	1959	1845	114	3.00	1.80	0.29	4.30
EIC_01D0	ESRL	Easter Island, Chile	27.15°S	109.45°W	50	288	288	0	7.50	0.03	0.87	0.93
ETL_06C0_14LST	EC	East Trout Lake, Saskatchewan, Canada	54.35°N	104.98°W	492	1865	1857	8	3.00	0.57	0.13	2.01
FSD_06C0_14LST	EC	Fraserdale, Canada	49.88°N	81.57°W	210	3605	3579	26	3.00	0.56	0.19	2.57
GMI_01D0	ESRL	Mariana Islands, Guam	13.43°N	144.78°E	3	723	723	0	1.50	0.37	−0.03	0.90
HBA_01D0	ESRL	Halley Station, Antarctica, United Kingdom	75.58°S	26.50°W	30	486	486	0	0.75	0.15	−0.02	0.20
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (μmol mol^{-1})**	**Innov. χ^2**	**Bias (μmol mol^{-1})**	**SE (μmol mol^{-1})**
HPB_01D0	ESRL	Hohenpeissenberg, Germany	47.80°N	11.01°E	985	210	206	4	7.50	0.74	2.70	5.83
HUN_01D0	ESRL	Hegyhatsal, Hungary	46.95°N	16.65°E	248	510	509	1	7.50	0.35	−0.12	4.37
ICE_01D0	ESRL	Storhofdi, Vestmannaeyjar, Iceland	63.40°N	20.29°W	118	486	482	4	1.50	0.55	−0.02	1.07
KEY_01D0	ESRL	Key Biscayne, Florida, United States	25.67°N	80.16°W	3	347	346	1	2.50	0.26	0.20	1.40
KUM_01D0	ESRL	Cape Kumukahi, Hawaii, United States	19.52°N	154.82°W	3	638	638	0	1.50	0.34	−0.02	0.95
KZD_01D0	ESRL	Sary Taukum, Kazakhstan	44.06°N	76.82°E	601	441	438	3	2.50	0.63	0.49	3.21
KZM_01D0	ESRL	Plateau Assy, Kazakhstan	43.25°N	77.88°E	2519	393	388	5	2.50	0.94	0.55	2.28
LEF_01C3_14LST	ESRL	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	3489	3412	77	3.00	0.87	0.25	2.85
LEF_01P0	ESRL	Park Falls, Wisconsin, United States	45.95°N	90.27°W	472	1350	1296	54	3.00	1.28	0.39	3.29
LLB_06C0_14LST	EC	Lac La Biche, Alberta, Canada	54.95°N	112.45°W	540	1225	1172	53	3.00	1.66	0.19	3.76
LMP_01D0	ESRL	Lampedusa, Italy	35.52°N	12.62°E	45	200	193	7	1.50	1.38	0.64	1.45
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (μmol mol^{-1})**	**Innov. χ^2**	**Bias (μmol mol^{-1})**	**SE (μmol mol^{-1})**
MAA_02D0	CSIRO	Mawson Station, Antarctica, Australia	67.62°S	62.87°E	32	206	206	0	0.75	0.26	−0.17	0.26
MHD_01D0	ESRL	Mace Head, County Galway, Ireland	53.33°N	9.90°W	5	407	407	0	2.50	0.18	0.09	1.06

		Galway, Ireland										
MID_01D0	ESRL	Sand Island, Midway, United States	28.21°N	177.38°W	4	497	496	1	1.50	0.61	0.52	1.12
MKN_01D0	ESRL	Mt. Kenya, Kenya	0.05°S	37.30°E	3897	134	134	0	2.50	1.12	1.85	1.89
MLO_01C0_02LST	ESRL	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	3342	3342	0	0.75	0.76	0.18	0.63
MLO_01D0	ESRL	Mauna Loa, Hawaii, United States	19.54°N	155.58°W	3397	566	566	0	1.50	0.22	0.13	0.68
MQA_02D0	CSIRO	Macquarie Island, Australia	54.48°S	158.97°E	12	265	265	0	0.75	0.28	0.05	0.33
NMB_01D0	ESRL	Gobabeb, Namibia	23.58°S	15.03°E	456	161	161	0	2.50	0.14	0.04	0.91
NWR_01D0	ESRL	Niwot Ridge, Colorado, United States	40.05°N	105.58°W	3523	495	491	4	1.50	0.66	0.40	1.38
NWR_03C0_02LST	NCAR	Niwot Ridge, Colorado, United States	40.05°N	105.58°W	3523	1550	1549	1	3.00	0.25	−0.36	1.32
OBN_01D0	ESRL	Obninsk, Russia	55.11°N	36.60°E	183	155	151	4	7.50	0.97	−1.16	8.26
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (µmol mol^{-1})**	**Innov. χ^2**	**Bias (µmol mol^{-1})**	**SE (µmol mol^{-1})**
OXK_01D0	ESRL	Ochsenkopf, Germany	50.03°N	11.80°E	1022	163	148	15	2.50	2.09	−0.31	3.62
PAL_01D0	ESRL	Pallas-Sammaltunturi, GAW Station, Finland	67.97°N	24.12°E	560	373	372	1	2.50	0.56	−0.12	1.89
POC_01D1	ESRL	Pacific Ocean, N/A	0.39°S	132.32°W	10	2043	2042	1	0.75	0.78	−0.01	0.63
PSA_01D0	ESRL	Palmer Station, Antarctica, United States	64.92°S	64.00°W	10	529	529	0	0.75	0.26	−0.09	0.27
PTA_01D0	ESRL	Point Arena, California, United States	38.95°N	123.74°W	17	381	381	0	7.50	0.43	−2.75	4.06
RPB_01D0	ESRL	Ragged Point, Barbados	13.17°N	59.43°W	45	499	499	0	1.50	0.47	0.21	1.09
SCT_01C3_14LST	ESRL	Beech Island, South Carolina, United States	33.41°N	81.83°W	115	675	667	8	3.00	1.00	0.08	2.94
SEY_01D0	ESRL	Mahe Island, Seychelles	4.67°S	55.17°E	3	481	481	0	0.75	1.07	0.06	0.72
SGP_01D0	ESRL	Southern Great Plains, Oklahoma, United States	36.80°N	97.50°W	314	407	388	19	2.50	1.42	−0.54	3.06
SGP_64C3_16LST	LBNL	Southern Great Plains, Oklahoma, United States	36.80°N	97.50°W	314	2625	2565	60	3.00	1.08	0.07	3.23
SHM_01D0	ESRL	Shemya Island, Alaska, United States	52.72°N	174.10°E	40	383	383	0	2.50	0.72	−0.03	2.01
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (µmol mol^{-1})**	**Innov. χ^2**	**Bias (µmol mol^{-1})**	**SE (µmol mol^{-1})**
SIS_02D0	CSIRO	Shetland Islands, Scotland	60.17°N	1.17°W	30	88	88	0	2.50	0.41	0.74	1.31
SMO_01C0_14LST	ESRL	Tutuila, American Samoa	14.25°S	170.56°W	42	3446	3446	0	0.75	0.49	0.16	0.50

Site code	Lab.	Location	Latitude	Longitude	Elev. (m ASL)	No. Obs. Avail.	No. Obs. Used	No. Obs. Rej.	r (μmol mol^{-1})	Innov. χ^2	Bias (μmol mol^{-1})	SE (μmol mol^{-1})
		Samoa										
SMO_01D0	ESRL	Tutuila, American Samoa	14.25°S	170.56°W	42	559	559	0	1.50	0.12	0.10	0.52
SNP_01C3_02LST	ESRL	Shenandoah National Park, United States	38.62°N	78.35°W	1008	716	696	20	3.00	1.13	−0.40	3.14
SPL_03C0_02LST	NCAR	Storm Peak Laboratory (Desert Research Institute), United States	40.45°N	106.73°W	3210	1543	1540	3	3.00	0.44	−0.52	1.75
SPO_01C0_14LST	ESRL	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	3882	3882	0	0.75	0.09	0.03	0.19
SPO_01D0	ESRL	South Pole, Antarctica, United States	89.98°S	24.80°W	2810	533	533	0	1.50	0.02	0.09	0.20
STM_01D0	ESRL	Ocean Station M, Norway	66.00°N	2.00°E	0	850	847	3	1.50	0.55	0.18	1.06
STR_01P0	ESRL	Sutro Tower, San Francisco, California, United States	37.76°N	122.45°W	254	642	632	10	3.00	0.64	0.08	2.55
SUM_01D0	ESRL	Summit, Greenland	72.58°N	38.48°W	3238	430	430	0	1.50	0.43	0.32	0.90
SYO_01D0	ESRL	Syowa Station, Antarctica, Japan	69.00°S	39.58°E	11	255	255	0	0.75	0.22	−0.14	0.24
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (μmol mol^{-1})**	**Innov. χ^2**	**Bias (μmol mol^{-1})**	**SE (μmol mol^{-1})**
TAP_01D0	ESRL	Tae-ahn Peninsula, Republic of Korea	36.73°N	126.13°E	20	383	383	0	7.50	0.40	0.60	4.33
TDF_01D0	ESRL	Tierra Del Fuego, Ushuaia, Argentina	54.87°S	68.48°W	20	192	192	0	0.75	0.44	−0.19	0.40
THD_01D0	ESRL	Trinidad Head, California, United States	41.05°N	124.15°W	107	369	324	45	2.50	3.25	−2.39	3.83
UTA_01D0	ESRL	Wendover, Utah, United States	39.90°N	113.72°W	1320	472	471	1	2.50	0.39	0.45	1.82
UUM_01D0	ESRL	Ulaan Uul, Mongolia	44.45°N	111.10°E	914	509	501	8	2.50	0.81	−0.21	2.60
WBI_01C3_14LST	ESRL	West Branch, Iowa, United States	41.72°N	91.35°W	242	1197	1134	63	3.00	1.81	0.37	3.77
WBI_01P0	ESRL	West Branch, Iowa, United States	41.72°N	91.35°W	242	862	795	67	3.00	2.07	0.22	4.18
WGC_01C3_14LST	ESRL	Walnut Grove, California, United States	38.27°N	121.49°W	0	1091	1021	70	3.00	1.68	−0.10	4.29
WGC_01P0	ESRL	Walnut Grove, California, United States	38.27°N	121.49°W	0	891	748	143	3.00	3.67	−0.67	6.51
WIS_01D0	ESRL	WIS Station, Negev Desert, Israel	31.13°N	34.88°E	400	540	537	3	2.50	0.61	0.03	1.94
WKT_01C3_14LST	ESRL	Moody, Texas, United States	31.31°N	97.33°W	251	2117	2100	17	3.00	0.65	0.14	2.29
Site code	**Lab.**	**Location**	**Latitude**	**Longitude**	**Elev. (m ASL)**	**No. Obs. Avail.**	**No. Obs. Used**	**No. Obs. Rej.**	**r (μmol mol^{-1})**	**Innov. χ^2**	**Bias (μmol mol^{-1})**	**SE (μmol mol^{-1})**
WKT_01P0	ESRL	Moody, Texas, United States	31.31°N	97.33°W	251	993	979	14	3.00	0.94	−0.10	2.85

		United States										
WLG_01D0	ESRL	Mt. Waliguan, Peoples Republic of China	36.29°N	100.90°E	3810	427	413	14	1.50	1.12	0.13	1.91
WSA_06C0_14LST	EC	Sable Island, Nova Scotia, Canada	49.93°N	60.02°W	5	2358	2259	99	3.00	1.17	0.17	4.12
ZEP_01D0	ESRL	Ny-Alesund, Svalbard, Norway and Sweden	78.90°N	11.88°E	475	568	566	2	1.50	0.96	0.55	1.22
All-site summary			-	-	-	85153	83910	1243	-	0.73	0.06	2.70

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 are the most common energy sources 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 2008, the world emissions of CO_2 from fossil fuel burning, cement manufacturing, and flaring reached 8.7 PgC yr^{-1} (one PgC=10^{15} grams of carbon) [Boden et al., 2011] and we estimate the global total emissions for 2009 and 2010 to be 8.6 PgC yr^{-1} and 9.1 PgC yr^{-1} respectively [Boden et al., 2011]. The 2010 figure represents a 47% increase over 1990 emissions. The North American (U.S.A, Canada, and Mexico) input of CO_2 to the atmosphere from fossil fuel burning was 1.8 PgC in 2008, representing 21% of the global total. North American emissions have remained nearly constant since 2000. On the other hand, emissions from developing economies such as the People's Republic of China have been increasing. The Department of Energy's 2011 International Energy Outlook has projected that the global total source will reach 9.1 PgC yr^{-1} in 2015 and 11.1 PgC yr^{-1} in 2030 [DOE].

Despite the recent economic slowdown, which affected developing countries starting in 2008, fossil fuel emissions in many parts of the world continue to increase.

In many flux estimation systems, including CarbonTracker, fossil fuel CO_2 emissions are specified. These imposed emissions are not optimized in the estimation framework. Thus, fossil fuel CO_2 emissions must be prescribed accurately in order to yield robust flux estimates for the land biosphere and oceans. Fossil fuel emissions estimates we use are available on an annually-integrated global and national basis, and this information needs to be gridded before being incorporated into CarbonTracker. The major uncertainty in this process is distributing the national-annual emissions spatially across a nation and temporally into monthly contributions. In CT2011, two different fossil fuel CO_2 emissions datasets were used to help assess the uncertainty in this mapping process. The legacy CarconTracker fossil fuel product ("Miller") has this year been augmented with the "ODIAC" [Oda and Maksyutov, 2011] emissions product. These two datasets share the same global and national emissions for each year, but differ in how those emissions are distributed spatially and temporally.

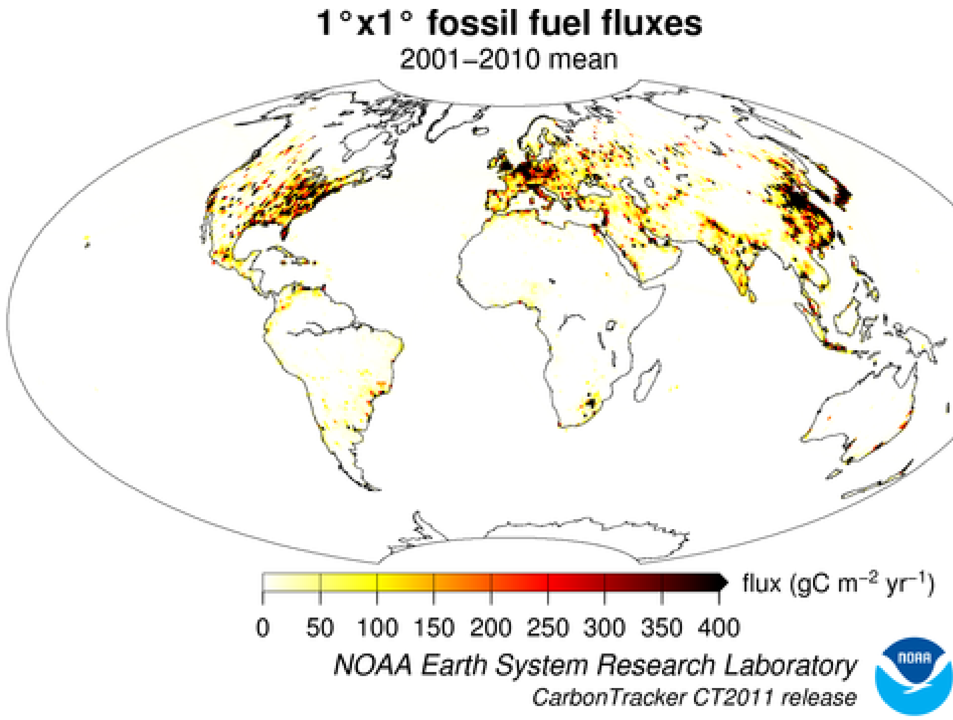


Figure 1. Spatial distribution of fossil fuel emissions. This is a spatial average of the Miller and ODIAC emissions inventories.

2. The "Miller" emissions dataset

- **Totals**
  The Miller fossil fuel emission inventory is derived from independent global total and spatially-resolved inventories. Annual global total fossil fuel CO_2 emissions are from the Carbon Dioxide Information and Analysis Center (CDIAC) [Boden et al. 2011] which extend through 2008. In order to extrapolate these fluxes to 2009 and 2010, we extrapolate using the percentage increase or decrease for each fuel type (solid, liquid, and gas) in each country from the 2011 BP Statistical Review of World Energy for 2009 and 2010.

- **Spatial Distribution**
  Miller fossil-fuel CO_2 fluxes are spatially distributed in two steps: First, the coarse-scale flux distribution country totals from Boden et al. [2011] are mapped onto a 1°x1° grid. Next, we distribute the country totals within countries according to the spatial patterns from the EDGAR v4.0 inventories [European Commission, 2009], which are annual estimates also at 1°x1° resolution. The CDIAC country-by-country totals, however, only sum to about 95% of the global total. We ascribe the difference to land regions according to the relative pattern of emissions over the globe.

- **Temporal Distribution**
  For North America between 30 and 60°N, the Miller system imposes a normalized, annually-invariant, seasonal cycle on emissions. This annual cycle is derived by extracting the first and second harmonics [Thoning et al, 1989] from the Blasing et al. [2005] analysis for the United States. The Blasing analysis has ~10% higher emissions in winter than in summer.

  For Eurasia, a set of seasonal emissions factors from EDGAR, distributed by emissions sector, is used to define fossil fuel seasonality. As in North America, this seasonality is imposed only from 30-60°N. The Eurasian seasonal amplitude is about 25%, significantly larger than that in North America, owing to the absence of a secondary summertime maximum due to air conditioning.

  See Box 1 for the resulting time series of fossil fuel emissions. In order to avoid discontinuities in the fossil fuel emissions between consecutive years, a spline curve that conserves annual totals fossil fuel emissions between consecutive years, a spline curve that conserves annual totals [Rasmussen 1991] is fit to seasonal emissions in each 1°x1° grid cell.

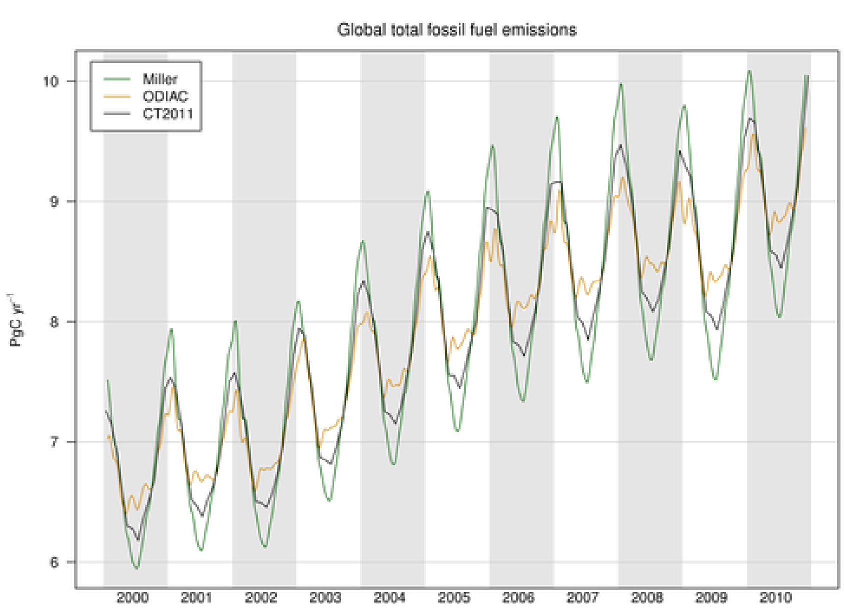
3. The "ODIAC" emissions dataset

- **Totals**
  The ODIAC fossil fuel emission inventory [Oda and Maksyutov, 2011] is also derived from independent global and country emission estimates from CDIAC, but from the previous year's estimates [Boden et al. 2010]. Annual country total fossil fuel CO_2 emissions from CDIAC which extend through 2007, were extrapolated to 2008, 2009 and 2010 using the BP Statistical Review of World Energy. The difference between the CDIAC global total and country-by-country totals were ascribed to the entire emissions fields. The same adjustment was done for 2009 and 2010 using preliminary 2009 and 2010 estimates by CDIAC.

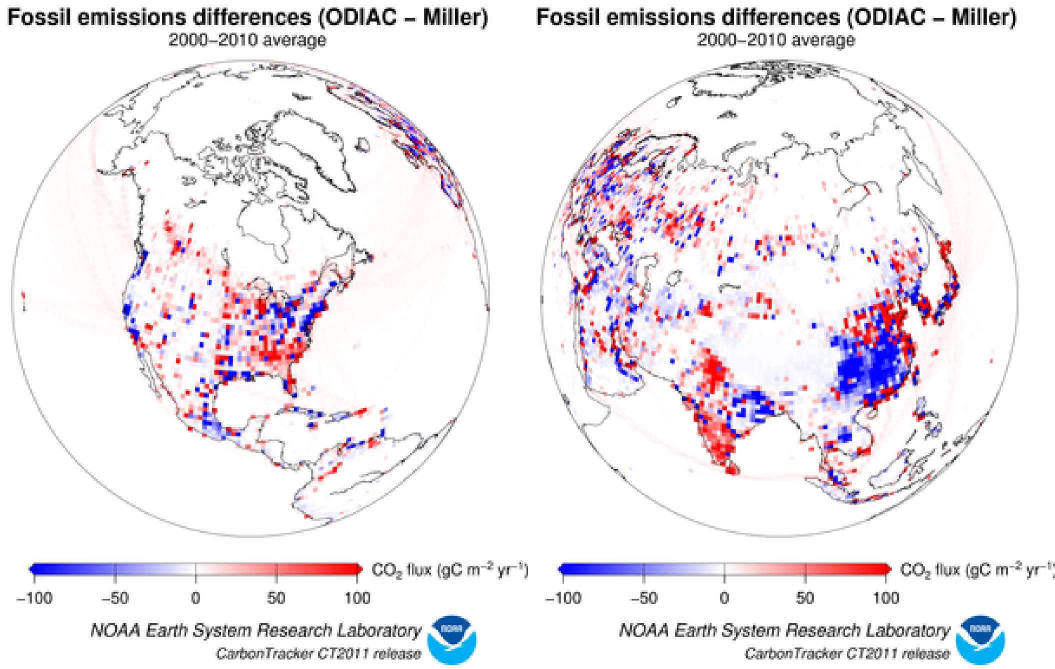
- **Spatial Distribution**
  ODIAC emissions are spatially distributed using many available "proxy data" that explain spatial extent of emissions according to emission types (emissions over land, gas flaring, aviation and marine bunker). Emissions over land were distributed in two steps: First, emissions attributable to power plants were mapped using geographical locations (latitude and longitude) provided by the global power plant data CARMA. Next, the remaining land emissions (i.e. land total minus power plant emissions) were distributed using nightlight imagery collected by U.S. Air Force Defense Meteorological Satellite Project (DMSP) satellites. Emissions from gas flaring were also mapped using nightlight imagery. Emissions from aviation were mapped using flight tracks adopted from UK AERO2k air emission inventory. It should be noted that currently, air traffic emissions are emitted at ground level within CarbonTracker. Emissions from marine bunker fuels are placed entirely in the ocean basins along shipping routes according to patterns from the EDGAR database.

- **Temporal Distribution**
  The CDIAC estimates used for mapping emissions in ODIAC only describe how much CO_2 was emitted in a given year. To present seasonal changes in emissions, we used the CDIAC 1°x1° monthly fossil fuel emission inventory [Andres et al. 2011]. The CDIAC monthly data utilizes the top 20 emitting countries' fuel (coal, oil and gas) consumption statistics available to estimate seasonal change in emissions. Monthly emission numbers at each pixel were divided by annual total and then a fraction to annual total was obtained. Monthly emissions in the ODIAC inventory were derived by multiplying this fraction by the emission in each grid cell.

Box 1. Comparison of the Miller and ODIAC global fossil fuel emissions estimates



Time series of global fossil fuel emissions. The Miller (green) and ODIAC (tan) estimates are each used by half of the eight inversions in the CT2011 suite, so the CT2011 (black) inventory is effectively an average of Miller and ODIAC. Note that fossil fuel emissions are not optimized in CarbonTracker.



Spatial differences in long-term mean fossil fuel emissions. between the two priors Note that both the Miller and ODIAC emissions inventories use the same country totals, but have different models for spatial distribution of that flux within countries.

- **Uncertainties**

The uncertainty attached to the global total source is of order 5% (2 sigma) until 2007 [Marland, 2008], but the uncertainties for individual regions of the world, and for sub-annual time periods are likely to be larger. Additional uncertainties are further introduced when the emissions are distributed in space and time. In the Miller dataset, the overall Eurasian seasonality is uncertain, but most likely a better representation than assuming no emission seasonality at all. Similarly, the use of the CDIAC monthly emission dataset for modeling seasonality introduces additional uncertainty in ODIAC. The additional uncertainty for the global total in the monthly CDIAC emission, which is solely due to the method for estimating seasonality, is reported as 6.4% [Andres et al. 2011]. As mentioned earlier, fossil fuel emissions are not optimized in the current CarbonTracker system, similar to many similar global carbon data analysis systems.

Spatial and temporal atmospheric CO_2 gradients arise from terrestrial biosphere and fossil-fuel sources. These gradients, which are interpreted by CarbonTracker, are difficult to attribute to one or the other cause. This is because the biospheric and anthropogenic sources are often co-located, especially in the temperate Northern Hemisphere.

Given that surface CO_2 flux due to biospheric activity and oceanic exchange is much more uncertain compared to fossil fuel emissions, CarbonTracker, like most current carbon dioxide data assimilation systems, does not optimize fossil fuel emissions. The contribution of CO_2 from fossil fuel burning to observed CO_2 mole fractions is considered known. However, for the first time in CarbonTracker, an effort is made to account for some aspects of fossil fuel uncertainty by using two different fossil fuel estimates as detailed above. From a technical point of view, extra land biosphere prior flux uncertainty is included in the system to represent the random errors in fossil fuel emissions. Eventually, fossil fuel emissions could be optimized within CarbonTracker, especially with the addition of $^{14}CO_2$ observations as constraints.

3. Further Reading

- CDIAC (Boden et al.) Annual Global and National fluxes
- DOE Energy Information Administration (EIA)
- BP Statistical Review of World Energy
- EDGAR Database
- CDIAC (Blasing et al.) Monthly USA fluxes
- L.A. Rasmussen "Piecewise Integral Splines of Low Degree", *Computers & Geosciences*, 17(9) pp 1255-1263, 1991
- Thoning et al. (1989) Atmospheric carbon dioxide at Mauna Loa Observatory. 2. Analysis of the NOAA GMCC data, 1974-85. *Journal of Geophysical Research* 94(D6), 8549-65.
- Marland, G. (2008), Uncertainties in Accounting for CO_2 from Fossil Fuels, *Journal of Industrial Ecology*, 12(2), 136-139.
- Boden, T.A., G. Marland, and R.J. Andres. 2011. Global, Regional, and National Fossil-Fuel CO_2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/00001_V2011
- CDIAC (Boden et al.) Preliminary 2009 and 2010 Global and National Estimates by Extrapolation
- The Center for Global Development, CARbon Monitoring Action (CARMA) power plant database
- DMSP satellite nightlight data
- Centre for Air Transport and the Environment (CATE), AERO2k aviation emissions inventory
- Marland, G. (2008), Uncertainties in Accounting for CO_2 from Fossil Fuels, *Journal of Industrial Ecology*, 12(2), 136-139.
- Andres et al. (2011) Monthly, global emissions of carbon dioxide from fossil fuel consumption. *Tellus B*, 63:309-327. doi: 10.1111/j.1600-0889.2011.00530.x.
- CDIAC (Andres et al.) Monthly Fossil-Fuel CO_2 emissions
- Oda, T. and Maksyutov, S. (2011) 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, doi:10.5194/acp-11-543-2011.
- European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). (2009) Emission Database for Global Atmospheric Research (EDGAR), release version 4.0

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 transport in the atmosphere: storm systems, cloud complexes, and weather of all sorts cause winds that transport CO_2 around the world. As a result, local events like fires, forest growth, and ocean upwelling can have impacts at remote locations. 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. Since CO_2 does not decay or react in the lower atmosphere, the influence of emissions and uptake in locations such as North America and Europe are ultimately seen in our measurements even at 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. To represent the atmospheric transport, we use the Transport Model 5 (TM5). This is a community-supported model whose development is shared among many scientific groups with different areas of expertise. TM5 is used for many applications other than CarbonTracker, including forecasting air-quality, studying the dispersion of aerosols in the tropics, tracking biomass burning plumes, and predicting pollution levels that future generations might have to deal with.

2. Detailed Description

TM5 is a global model with two-way nested grids; regions for which high-resolution simulations are desired can be nested in a coarser grid spanning the global domain. The advantage to this approach is that transport simulations can be performed with a regional focus without the need for boundary conditions from other models. Further, this approach allows measurements outside the "zoom" domain to constrain regional fluxes in the data assimilation, and ensures that regional estimates are consistent with global constraints. TM5 is based on the predecessor model 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 al., 2003). 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 separately simulates advection, convection (deep and shallow), and vertical diffusion in the planetary boundary layer and free troposphere.

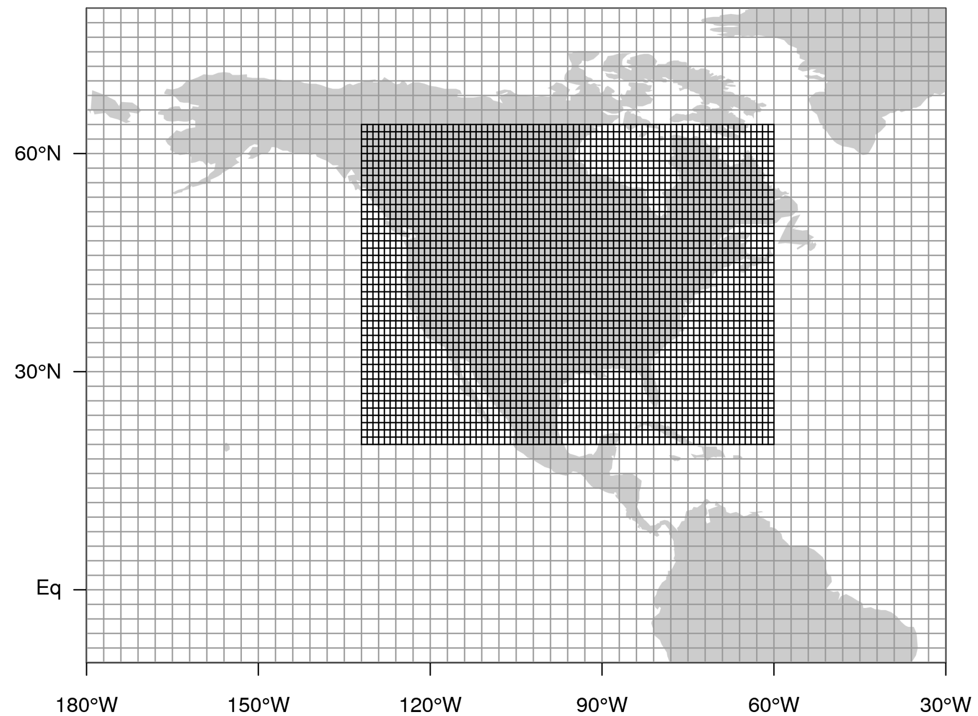


Figure 1. TM5 grids used in CarbonTracker. Figure shows the 1°x 1° nested regional grid over North America and a portion of the global 3°x 2° grid.

The winds which drive TM5 come from the European Center for Medium range Weather Forecast (ECMWF) operational forecast model. This "parent" model currently runs with ~25 km horizontal resolution and 60 layers in the vertical prior to 2006 (and 91 layers layers in the vertical from 2006 onwards). The carbon dioxide levels predicted by CarbonTracker do not feed back onto these predictions of winds.

For use in TM5, the ECMWF meteorological data are preprocessed into coarser grids. In CarbonTracker, TM5 is run at a global 3°x 2° resolution with a nested regional grid over North America at 1° x 1° resolution (Figure 1). 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 34 hybrid sigma-pressure levels (from 2006 onwards; 25 levels for 2000-2005), unevenly spaced with more levels 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

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

- 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, heading, speed, and other factors, 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" [Kaminski et al., 2001]. We chose an approach in which the ocean is divided up into 30 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 below. Figure 1 shows ecoregions for North America (click here for global land ecoregions). 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. Further details on ecoregions can be found here.

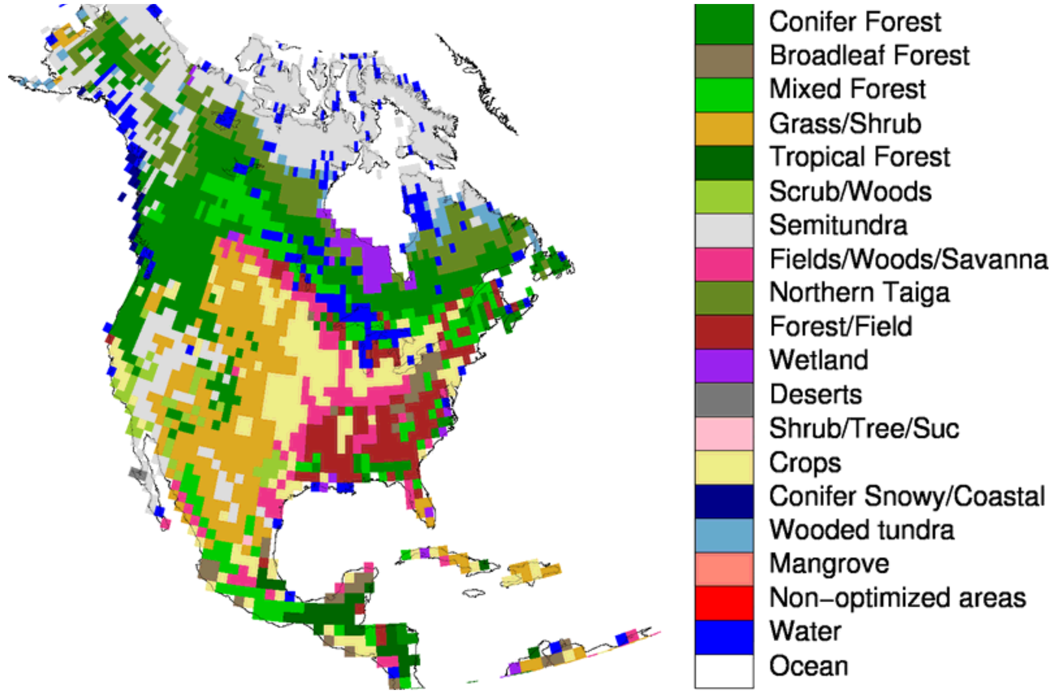


Fig 1. CarbonTracker ecoregions in North America

Theoretically, this approach leads to a total number of 11*19+30=239 optimizable scaling factors λ each week, but the actual number is 156 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 desert. 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 1° x 1° 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 1° x 1° 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 similar (from our perspective) 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%

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 1° x 1° pixel of our domain was assigned one of the categories above bases on the Olson category that was most prevalent in the 0.5° x 0.5° 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 its 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. If the relationship between a parameter deviation and its modeled observational impact is statistically significant, then that relationship is used to modify parameters. Otherwise, the relationship is assumed to be spurious noise due to the numerical approximation of the covariance matrix by the limited ensemble. We accept relationships that reach 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.

3 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 correlations between 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 spurious compensating source/sink pairs ("dipoles") to occur in the tropics.

In our standard assimilation, the chosen standard deviation is 80% on land parameters. 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.

3.1 Multiple prior models

In Bayesian estimation systems like CarbonTracker, there is a potential for bias from a flux prior to propagate through the inversion system to the final result. It is difficult to quantify this effect, and as a result it is generally considered a precondition that flux priors be unbiased. We cannot guarantee this for any of our fluxes, be they the prior estimates for terrestrial or oceanic exchange, or the presumed wildfire and fossil fuel emissions. In order to explicitly quantify the impact of prior bias on our solution, in ~~CT2011~~ CT2011_oi we present the result of a multi-model prior suite of inversions. We have used two terrestrial flux priors, ~~two air-sea exchange priors~~ one air-sea exchange prior, and two estimates of imposed fossil fuel emissions in a three-way factorial design experiment. This has resulted in ~~eight~~ four individual inversions, each using a unique combination of priors and conducted independently according to the methods described above. We present as a final result the mean flux across this suite of inversions and the atmospheric CO_2 distribution resulting from applying these mean fluxes to our atmospheric transport model. Each of the priors is described in detail on its corresponding documentation page (fossil, land, ocean).

Notice: CT2011_oi does not use the "climatological" ocean prior. After we released CarbonTracker 2011, a significant bug was discovered in our atmospheric transport model. We have corrected the bug and are releasing revised results under the release name "CT2011_oi". One consequence of this problem is that the four inversions using the climatological ocean flux prior were faulty. They have been removed from the inversion suite in CT2011_oi. Use of the original CT2011 results is strongly discouraged. Details can be found at this link.

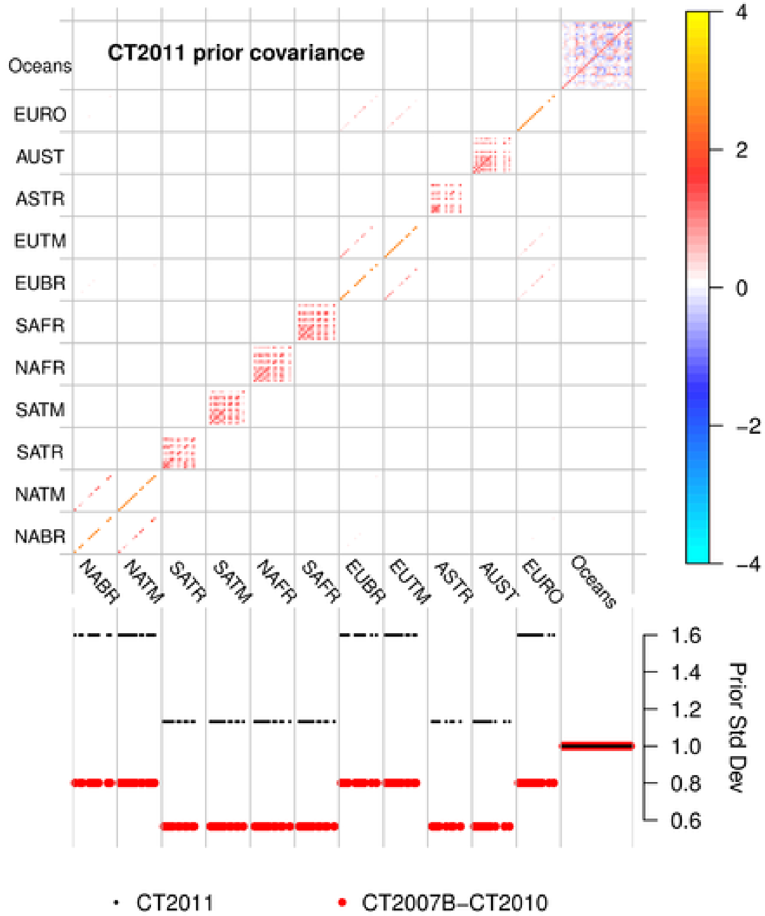


Fig 2. CT2011_oi prior covariance structure. The prior covariance matrix (top panel) and the square root of diagonal members of this matrix (bottom panel). Covariance matrix quantities are dimensionless squared scaling factors, and the bottom panel is the square root of this. TransCom land regions form the first 11 large divisions on the axes here. As described above, each of those regions contains 19 potential ecosystems. Correlations between similar ecosystems in proximate Transcom regions are visible in North America (e.g. NABR and NATM, the boreal and temperate North American regions) and Eurasia. Within tropical Transcom regions, however, differing ecosystems are assigned a non-zero prior covariance, which is visible here as red block-like structures on the diagonal within, for example, the South America Tropical (SATR) Transcom region. Ocean regions have a more complicated covariance structure that depends on which prior is used; the structure shown here is that of the ocean inversion flux prior. The lower panel of this diagram compares the on-diagonal elements of the prior covariance matrix by plotting their square roots. The resulting standard deviations are directly comparable to the percentages discussed in section 3 above; 0.8 is equivalent to 80%. The retuning of the covariance matrix for CT2011_oi's multiple-prior simulation is made evident by also showing these values from previous CarbonTracker releases in red.

3.2 Posterior Uncertainties in CarbonTracker

The formal "internal" error estimates produced by CarbonTracker are unrealistically large. This is largely a result of the relatively short assimilation window in CarbonTracker, along with a dynamical model that introduces a fresh prior covariance matrix with every new week entering the assimilation window. This five-week window effectively inhibits the formation of anticorrelations ("dipoles") in flux estimates, and does little to reduce the confidence interval on prior fluxes.

The temporal truncation in CarbonTracker imposed by its five-week assimilation window tends to yield regional flux estimates that are largely uncorrelated with those from other regions. A consequence of this feature is that uncertainties in CarbonTracker tend to increase as larger regions are considered; regional errors mostly just add in quadrature without any cancellation from dipole anticorrelation. Whereas many inversions yield smaller errors as the spatial extent of the region being considered increases, CarbonTracker acts in the opposite fashion. This is perhaps most obvious in the estimate of CarbonTracker's global annual surface flux of carbon dioxide. While CT2011_oi estimates a one-sigma error of more than 6 $PgCyr^{-1}$ on its global flux, this quantity is in actuality much more well-constrained. This is evident from CarbonTracker's excellent agreement with observational estimates of atmospheric growth rate.

In CT2011_oi, error estimates are about a factor of two larger than in previous releases, mainly due to the retuning of the land prior covariance discussed above. However, uncertainties presented for CT2011_oi take into account not only the "internal" flux uncertainty generated by a single inversion, but also the across-model "external" uncertainty representing the spread of the inversion models due to the choice of prior flux.

4. Further Reading

- Whitaker and Hamill, 2002 paper
- Peters et al., 2005 paper
- Olson ecosystem types, data
- Kaminski et al., *Journal of Geophysical Research*, Vol. 106, No. D5, PP. 4703-4715, 2001 doi:10.1029/2000JD900581

Ecoregions in CarbonTracker [goto top]

1. What are ecoregions?

Ecoregions are the actual scale on which CarbonTracker performs its optimization over the land. Ecoregions are meant to represent large expanses of land within a given continent having similar ecosystem types, and are used to divide continents into smaller pieces for analysis. The ecosystem types use in CarbonTracker are derived from the Olson [1992] vegetation classification (Table 1, Figure 1).

We define an ecoregion as an ecosystem type within a given Transcom land region. There are 11 such Transcom land regions (Figure 2), so there are 11*19 = 209 possible ecoregions. However, not all ecosystem types are present in all Transcom regions, and the actual number of land ecoregions ends up being 126.

Note on "Semitundra": this is a potentially misleading shorthand abbreviation for a collection of ecosystems comprising semi-desert, shrubs, steppe, and polar+alpine tundra. The "Semitundra" zones appearing in northern Africa where one expects to find the Sahara desert are not, of course, tundra environments. They are instead semi-desert zones.

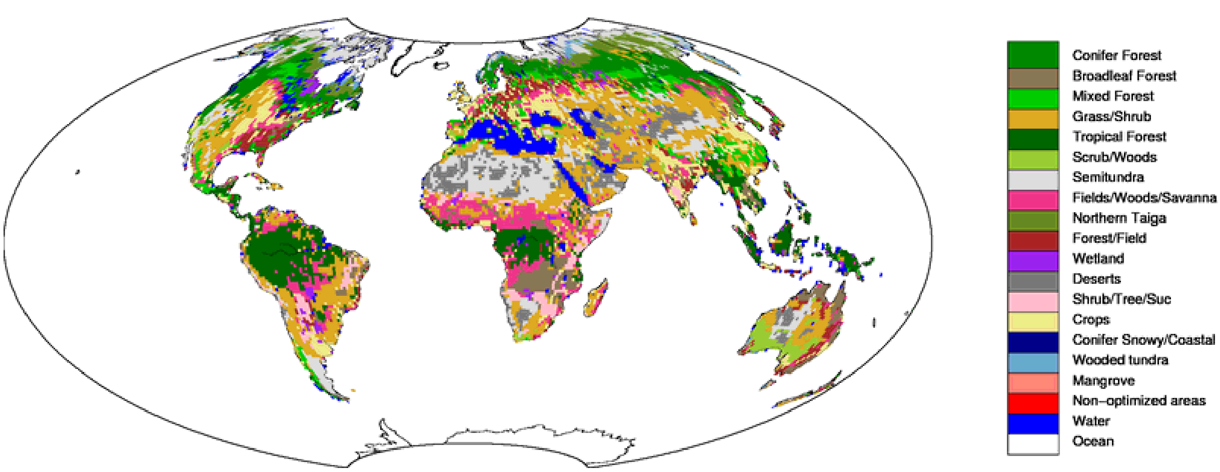


Figure 1. Global distribution of Olson ecosystem types.

Table 1. Ecosystem areas over the two Transcom regions covering North America.

Ecosystem Type	North American Boreal	North American Temperate

	Area (km2)	Percentage	Area (km2)	Percentage
Conifer Forest	2315376	22.9%	1607291	14.0%
Broadleaf Forest	–	–	269838	2.4%
Mixed Forest	592291	5.9%	930813	8.1%
Grass/Shrub	53082	0.5%	2515582	21.9%
Tropical Forest	–	–	58401	0.5%
Scrub/Woods	–	–	416520	3.6%
Semitundra	3396292	33.6%	866468	7.6%
Fields/Woods/Savanna	29243	0.3%	1020939	8.9%
Northern Taiga	1658773	16.4%	–	–
Forest/Field	61882	0.6%	1243174	10.8%
Wetland	322485	3.2%	66968	0.6%
Deserts	–	–	21934	0.2%
Shrub/Tree/Suc	–	–	11339	0.1%
Crops	–	–	1969912	17.2%
Conifer Snowy/Coastal	41440	0.4%	73437	0.6%
Wooded tundra	360388	3.6%	6643	0.1%
Mangrove	–	–	–	–
Non-optimized areas	–	–	–	–
Water	1269485	12.6%	384728	3.4%
Total	10100736	100.0%	11463986	100.0%

2. Why use ecoregions?

A fundamental challenge to atmospheric inversions like CarbonTracker is that there are not enough observations to directly constrain fluxes at all times and in all places. It is therefore necessary to find a way to reduce the number of unknowns being estimated. Strategies to reduce the number of unknowns in problems like this one generally impose information from external sources. In CarbonTracker, we reduce the problem size both by estimating fluxes at the ecoregion scale, and by using a terrestrial biological model to give a first guess flux from the ecoregion. The model is also used to give the spatial and temporal distribution of CO_2 flux within a region and week.

2. Ecosystems within Transcom regions

Each Transcom land region (Figure 2) can contain up to 19 ecoregions.

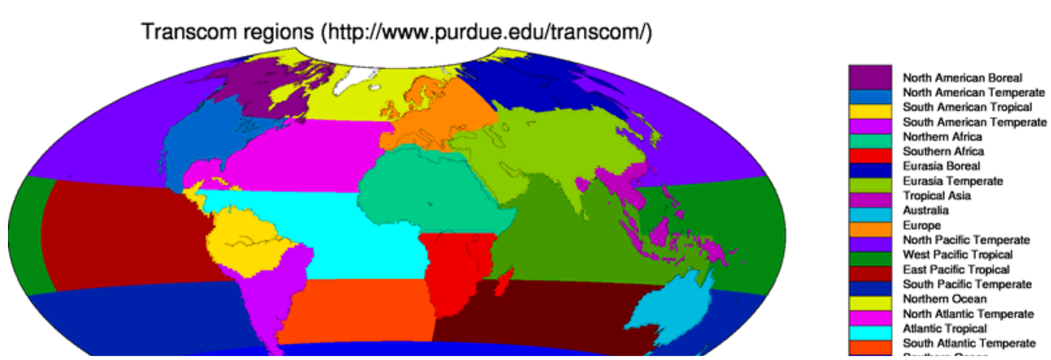


Figure 2. The 11 land regions and 11 ocean regions of the [Transcom project](#).

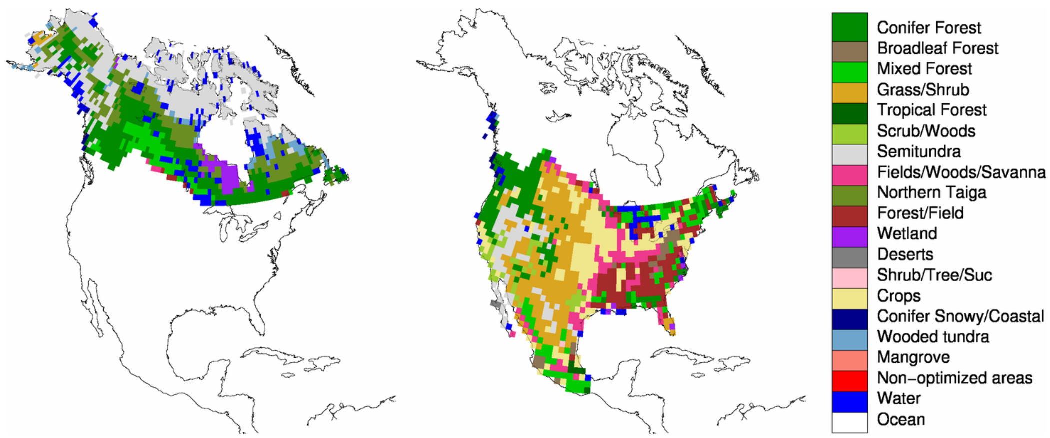


Figure 3. Ecoregions within the North American Boreal (left) and North American Temperate (right) Transcom regions.

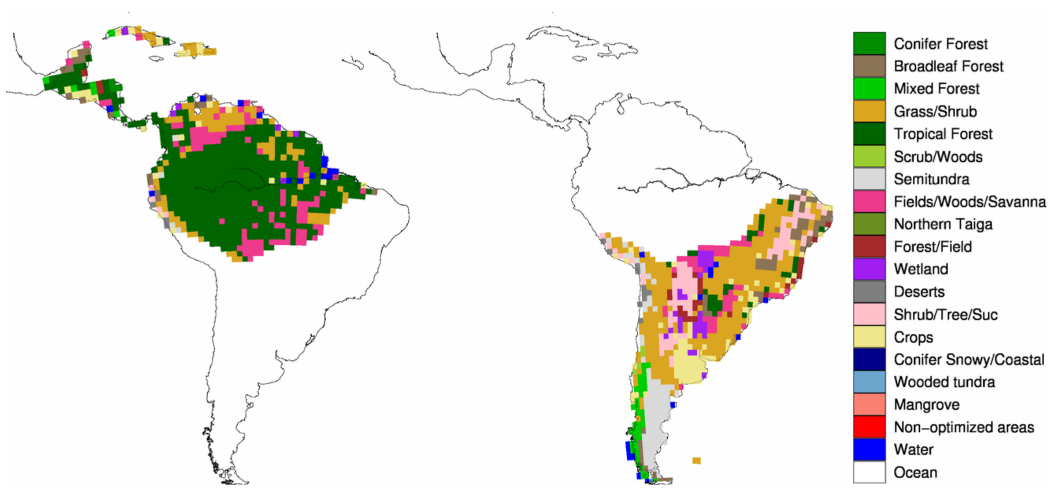


Figure 4. Ecoregions within the South American Tropical (left) and South American Temperate (right) Transcom regions.

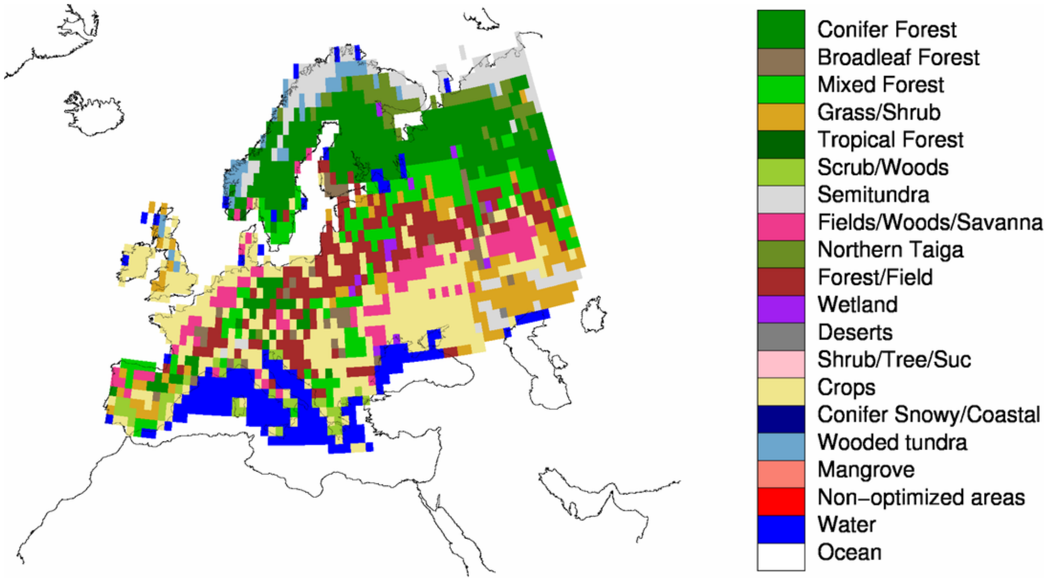


Figure 5. Ecoregions within the Europe Transcom region.

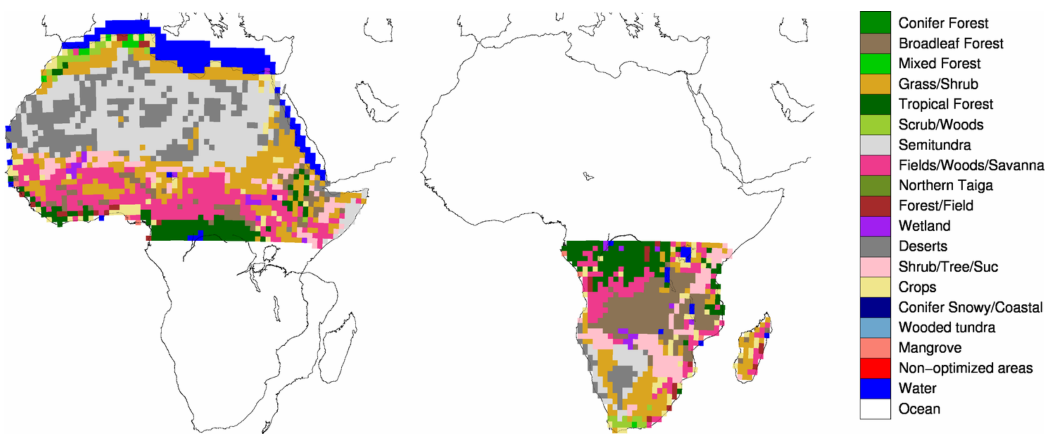


Figure 5. Ecoregions within the Northern Africa (left) and Southern Africa (right) Transcom regions.

3. Further Reading

- Olson ecosystem types, data