Tellus
SERIES B
CHEMICAL
AND PHYSICAL
METEOROLOGY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

*Tellus (2010), 62B, 797–809*



TELLUS

# New technique to analyse global distributions of $CO_2$ concentrations and fluxes from non-processed observational data

*By* T. MAKI[1]*, M. IKEGAMI[2], T. FUJITA[2], T. HIRAHARA[2], K. YAMADA[2], K. MORI[3], A. TAKEUCHI[4], Y. TSUTSUMI[5], K. SUDA[2] and T. J. CONWAY[6], [1]*Atmospheric Environment and Applied Meteorology Research Department, Meteorological Research Institute, Japan Meteorological Agency, Japan;* [2]*Global Environment and Marine Department, Japan Meteorological Agency, Japan;* [3]*Kobe Marine Observatory, Japan Meteorological Agency, Japan;* [4]*Meteorological Satellite Center, Japan Meteorological Agency, Japan;* [5]*Mito Local Meteorological Observatory, Japan Meteorological Agency, Japan;* [6]*Global Monitoring Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, USA*

(Manuscript received 25 November 2009; in final form 22 June 2010)

ABSTRACT

We have developed a new observational screening technique for inverse model. This technique was applied to our transport models with re-analysed meteorological data and the inverse model to estimate the global distribution of $CO_2$ concentrations and fluxes. During the 1990s, we estimated a total $CO_2$ uptake by the biosphere of 1.4–1.5 PgC $yr^{-1}$ and a total $CO_2$ uptake by the oceans of 1.7–1.8 PgC $yr^{-1}$. The uncertainty of global $CO_2$ flux estimation is about 0.3 PgC $yr^{-1}$. We also obtained monthly surface $CO_2$ concentrations in the marine boundary layer to precisions of 0.5–1.0 ppm. To utilize non-processed (statistical monthly mean) observational data in our analysis, we developed a quality control procedure for such observational data including a repetition of inversion. This technique is suitable for other inversion setups. Observational data by ships were placed into grids and used in our analysis to add to the available data from fixed stations. The estimated global distributions are updated and extended every year.

## 1. Introduction

Of the greenhouse gases, $CO_2$ contributes most to global warming. However, the distribution of its sources and sinks is not fully understood (IPCC, 2007). The World Meteorological Organization (WMO) has launched the Global Atmosphere Watch (GAW) programme, which recommends, in its strategic plan for 2008–2015, that observational data be integrated by numerical models and data assimilation techniques according to the International Global Atmospheric Chemistry Observations (IGACO) strategy. Such integrated products are important and effective tools for environmental monitoring. Following this strategy, the Japan Meteorological Agency (JMA) has provided an integrated product of the distribution of $CO_2$ concentrations since 2009 by putting non-processed (statistical monthly mean) observational data from cooperative laboratories into our transport model and the inverse model, which is one of the most useful tools for this purpose (Rayner et al., 1999; Gurney et al., 2002; Rödenbeck et al., 2003; Gurney et al., 2004; Baker et al., 2006; Stephens et al., 2007; Gurney et al., 2008). Although many studies use processed observational data in the inverse model, we used non-processed observational data in the analysis to estimate $CO_2$ fluxes with a high spatial and temporal resolution because there must be some signals that are contained only in non-processed data. However, some observational data are affected by local fluxes and such data may easily give adverse impacts on analysis results; the quality of the estimation depends on how to screen non-processed data. Therefore, we have developed a new data screening technique for estimating $CO_2$ distributions. The important feature of our analysis is that we only used data from selected observational site that were screened by utilizing results from the inverse model. This method is simple but rational in that we can select observational data that match our inversion setup (number of regions, pre-subtracted and prior fluxes, uncertainty of fluxes and distribution of observational sites). If certain observational data are used only in a specific phase in the inversion, we can do so by determining flexibly which data to be used in one phase but not in another. Michalak et al. (2005)

*Corresponding author.
e-mail: tmaki@mri-jma.go.jp
DOI: 10.1111/j.1600-0889.2010.00488.x

showed an important approach in surface flux inversions. In this method, observational errors were estimated using a Maximum Likelihood approach. However, this method did not completely reject observational data, which were affected by local fluxes and might cause impacts on analysis results.

## 2. Analysis method

### 2.1. Inversion method

We utilized a Bayesian Synthesis Inversion method (Tarantola, 1987; Enting, 2002) to estimate surface $CO_2$ fluxes from observational data, in combination with an atmospheric transport model and prior fluxes statistically determined for 22 regions. Our analysis method is based on TransCom 3, which is described by Baker et al. (2006). The methods and parameters for our analysis are summarized here. We assumed that observed $CO_2$ concentrations be determined linearly from regional fluxes through a matrix of transport coefficients, as shown in eq. (1), where $y_i$, and $x_j$ are observed $CO_2$ concentrations and regional $CO_2$ fluxes, respectively, and $a_{ij}$ are coefficients representing the contributions of regional fluxes to observations which are calculated by the transport model.

In general, meaningless fluxes may be calculated when eq. (1) is solved simply to find $CO_2$ fluxes that minimize the differences between observations and model simulations. To avoid this disadvantage, we used the following techniques. First, prior fluxes ($x_p$) and their uncertainties ($C_x$) were introduced to determine an $x$ that minimized the cost function $S(x)$ in eq. (2), which calculated more suitable regional fluxes using prior information as shown in the second term of the right-hand side in eq. (2). Secondly, the transport model results from pre-subtracted fluxes (emission from fossil fuel burning and biospheric and oceanic exchanges) were subtracted from observations and then residual concentrations were inversed using eq. (1). In eqs. (2) and (3), $C_y$ is the uncertainty of observed $CO_2$ concentrations and the matrix $A$ ($a_{ij}$) contains the result of transport model from all unit fluxes and $A^{\mathrm{T}}$ means its transposed (adjoint) matrix. The posterior regional flux estimate $x$ may

$$\begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & . & . & . & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \quad (1)$$

$$S(x) = (Ax - y)^T C_y^{-1}(Ax - y) + (x - x_p)^T C_x^{-1}(x - x_p) \quad (2)$$

$$x = x_p + \left(A^{\mathrm{T}} C_y^{-1} A + C_x^{-1}\right)^{-1} A^{\mathrm{T}} C_y^{-1}(y - Ax_p) \quad (3)$$

be solved for analytically as eq. (3). From the calculated posterior $CO_2$ fluxes and subsequent simulations with the transport model, $CO_2$ concentrations were obtained globally in three-dimensional grids.

*2.1.1. Prior fluxes and their uncertainties*. Surface fluxes were estimated for 22 geographical regions, as for TransCom 3 (Gurney et al., 2000), for the period from 1985 to 2007. The transport model was run for 3 years with a pre-determined pattern of $CO_2$ fluxes in the re-analysed wind fields (JRA-25, Onogi et al., 2007) to obtain pre-determined regional flux patterns for different regions and months of the year in a way consistent with TransCom 3 (Baker et al., 2006). We also adopted prior source fluxes and their uncertainties as done in the previous study (Baker et al., 2006).

To obtain realistic flux estimates, pre-subtracted fluxes were used as for TransCom 3. The spatial pattern of emissions from fossil fuel burning in 1990 (Andres et al., 1996) and 1995 (Brenkert, 1998), CASA biosphere fluxes (Randerson et al., 1997) and oceanic exchanges (Takahashi et al., 1999) were treated as known fluxes and the contributions of these known fluxes were subtracted from the observed concentrations before calculating the minimal cost function. Recent $CO_2$ emissions from fossil fuel burning were estimated from annual global statistics from 1980 to 2007 (Marland et al., 2007). We used the values for 2007 for the years thereafter.

*2.1.2. Observational data and their uncertainties*. Previous studies (Baker et al., 2006; Gurney et al., 2008) used processed data from GLOBALVIEW (NOAA, 2007), which were smoothed, interpolated and extrapolated. However, such data may not retain signals that can only be found in non-processed observational data. In this study, we used monthly mean observational data submitted to the WDCGG from different laboratories for the inversion (Rödenbeck et al., 2006). This is an important feature of our analysis. To put non-processed observational data into the inversion, the data were subjected to quality control and their uncertainties were estimated. In the first step of quality control, only data traceable to the WMO standard scale were selected (Zhao and Tans, 2006). In the second step, we chose the longest one from plural records in a single location. Next, missing data were filled by interpolation or extrapolation (Nakazawa et al., 1991; WMO, 2000) to create data sets representing average seasonal cycle. Uncertainties of the data were defined as the standard deviation of the differences between the observed and the smoothed data as done by many previous studies. Data in periods of no observation were filled with the data created earlier, but were given a large uncertainty (150 ppm) to avoid affecting analysis results. The data were subsequently screened in the following steps (Fig. 1):

1. All available observational data were put into the inversion model to obtain global estimates of concentrations.

2. Monthly observations with a larger mismatch (analysis–observation) than the threshold (1 sigma for the first cycle or three sigmas for the subsequent cycles. The sigma is the square root of the mean squared value of the distance of each analysis

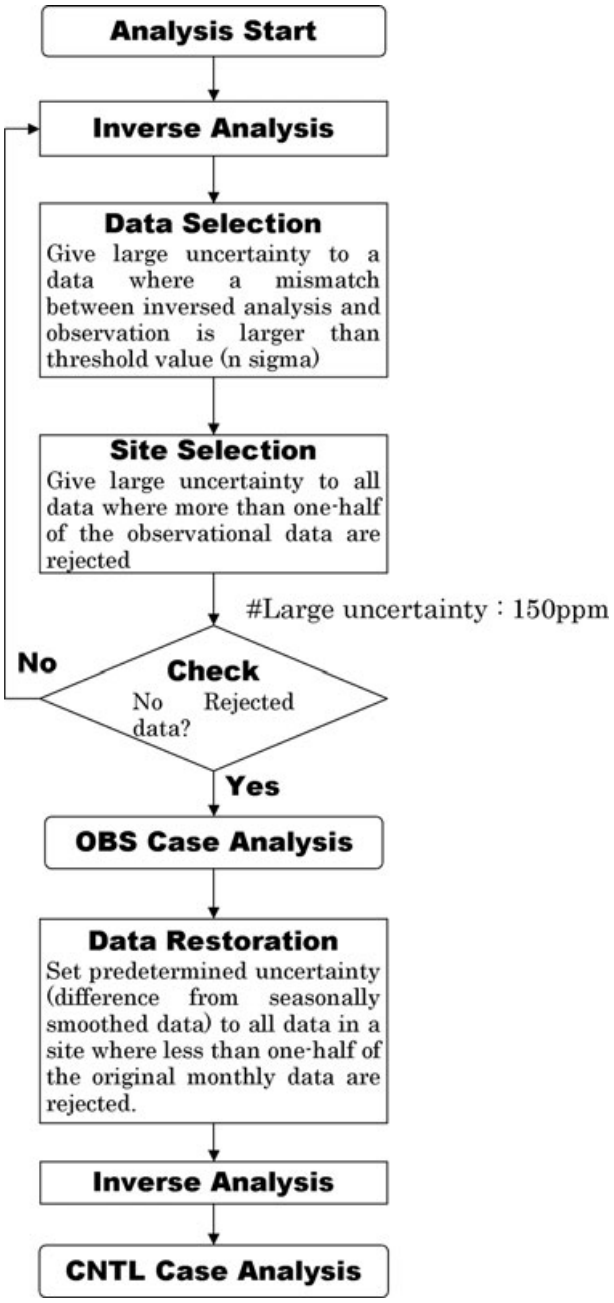


*Fig. 1.* A flow-chart for observational data and site selection procedure.

from the observation for all observational sites) were rejected (given a large uncertainty (150 ppm)): data selection.

3. To reduce computational cost, large uncertainties were given to all observational data in the sites from which more than a half of observational data were rejected (site selection).

4. If any data were rejected in Step 2, the procedure from Step 1 was repeated.

5. The inversion results (concentrations) so far were treated as those in the OBS case.

6. Predetermined uncertainties, defined as the standard deviations from the seasonally smoothed data, were given to the sites from which less than a half of observational data were rejected (data restoration in the selected sites).

7. As the final step, the inversion was performed once again and the resulting concentrations were treated as those in the CNTL case.

In general, we repeated three to four times of Steps 1–4 to finish the analysis. This process leads to the rejection of observational data that differ from the concentration estimated by the inverse model based on a majority of the observational data. It is emphasized that the rejected data are not represented in the sparse horizontal and spatial resolution of the model. The observational sites are distributed as shown in Fig. 2, in which colours indicate the rates of monthly values selected before Step 5. In general, the rates are large at oceanic sites and in the southern hemisphere, but small at land sites and in the Northern Hemisphere. In the European area, there are many sites where data selection rate is extremely small. In this area, many data measured in summer and winter were rejected. This is likely due to a large spatial flux variability from local anthropogenic or biospheric fluxes. Therefore, a larger number of geographical regions are required in model calculations for a more detailed analysis.

In addition to the data from fixed stations on land, we used observational data from ships in the western Pacific; these data were placed into grids by averaging over $5^\circ \times 5^\circ$ square boxes and then over a month to increase the horizontal resolution of the analysis and minimize the period of no observational data. The uncertainties of such data were fixed at 0.5 ppm, based on the uncertainties of neighbouring sites and the temporally uneven distribution of observational data.

As a result, data from as many as 146 sites/grids, of which 58 were from fixed land stations, 13 from flask aircraft observations at the upper troposphere and lower stratosphere (Matsueda et al., 2002; Machida et al., 2008) and 75 from ship observations, were selected for use in the analysis. These numbers were larger than in previous studies. The important feature of this observational network is that we could select more than 90% of flask aircraft observations (The north–south circle coloured red from Japan to Australia in Fig. 2). This means that a combination of our transport model and analysed surface flux could represent $CO_2$ concentrations variability in lower troposphere. This may be supported by previous study (Stephens et al., 2007). The contributing institutions are listed in Table 1 and a time series of the numbers of sites/grids is shown in Fig. 3. There were about 70–80 sites/grids available in the 1990s, but only 20–40 sites/grids in the 1980s. In total, we introduced 21 738 monthly mean observations, of which 19 527 were selected in the CNTL case and 17 539 in the OBS case. We also tested to use GLOBALVIEW data experiment without data selection procedure shown earlier, which is referred to as the GV case, in which

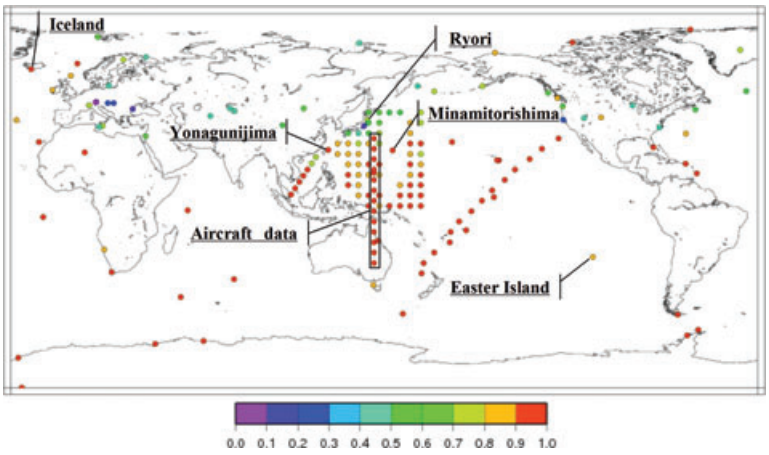


*Fig. 2.* Rates of selected data during the analysis.

*Table 1.* List of institutions that provided data for the global analysis

| Country | Institution | Number of sites | Reference |
|---|---|---|---|
| Argentina | Direcion Nacional del Antartico- Istituto Antartico Argentino | 1 | Ciattaglia et al. (1999) |
| Australia | Bureau of Meteorology (BoM) | 1 | Francey et al. (1993) |
| Australia | Commonwealth Scientific and Industrial Research Organisation (CSIRO) | 6 | Francey et al. (2003) |
| Canada | Environment Canada (EC) | 2 | Higuchi et al. (2003) |
| Finland | Finnish Meteorological Institute | 1 | Eneroth et al. (2005) |
| Hungary | Hungarian Meteorological Service | 2 | Haszpra et al. (2001) |
| Italy | ENEA-Ricerca sul Sistema Elettrico S.p.A. (former CESI RICERCA) | 1 | Apadula et al. (2003) |
| Italy | Italian National Agency for New Technology, Energy and the Environment | 1 | Chamard et al. (2003) |
| Italy | International Center for Earth Sciences (ICES) | 1 | Ciattaglia et al. (1987) |
| Japan | Japan Meteorological Agency Meteorological Research Institute | 72 | Watanabe et al. (2000), Matsueda et al. (2002) |
| Russian Federation | Main Geophysical Observatory (MGO) | 5 | Paramonova et al. (2001) |
| South Africa | South African Weather Service | 1 | Brunke et al. (2004) |
| Spain | Izaña Atmospheric Research Center. Meteorological State Agency of Spain | 1 | Navascues and Rus (1991) |
| Sweden | Department of Applied Environmental Science, Stockholm University | 1 | Engardt et al. (1996) |
| United States of America | National Oceanic and Atmospheric Administration/Global Monitoring Division (NOAA/GMD) | 68 | Conway et al. (2007) |

18 632 data were used. The details are shown in Supporting Information.

*2.1.3. Transport model and forward calculations*. JMA's transport model (JMA-CDTM, Sasaki et al., 2003) was used for the analysis. This is an offline model based upon JMA's operational global spectral model (Sugi et al., 1990) to calculate distributions of passive tracers ($CO_2$, $^{222}Rn$, $SF_6$, etc). The model has a horizontal resolution of 2.5° and 32 vertical layers. It can simulate horizontal and vertical transport, vertical diffusion by turbulence (Mellor and Yamada 1974), cumulus convection (Kuo, 1974) and shallow convection with surface fluxes of given tracers. The model was utilized in several TransCom experiments and received good evaluations in terms of vertical mixing strength (Stephens et al., 2007). The vertical mixing coefficients of JMA-CDTM are relatively high compared with other transport models (Law et al., 2008; Patra et al., 2008). Data from the Japanese 25-year re-analysis (JRA-25, Onogi et al., 2007) and JMA Climate Data Assimilation System (JCDAS), following the same methodology as JRA-25, were used to calculate the transport of tracers. This high-quality and long-term meteorological

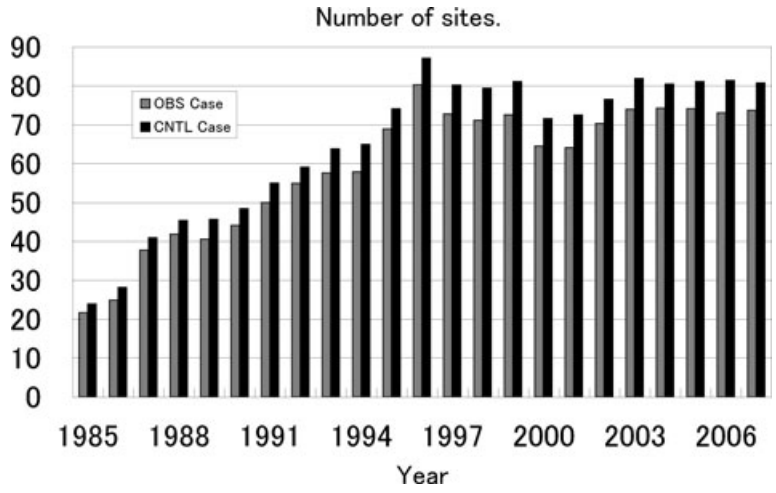


*Fig. 3.* Number of observational sites providing data for the analysis (from 1985 to 2007).

*Table 2.* Main features of the analysis method

| Analysis method | Bayesian synthesis inversion adopted in TransCom 3 (Baker et al., 2006) |
|---|---|
| Analysis period | From 1985 to 2007 |
| Time resolution | 1 Month |
| Horizontal resolution | 22 Regions (11 land and 11 ocean regions, as in TransCom 3) |
| Pre-subtracted fluxes | Fossil fuel burning (1990 and 1995), vegetation (CASA) and ocean (Takahashi et al., 1999) |
| Prior regional fluxes | Not applicable |
| Uncertainty of prior fluxes | Same as TransCom 3 |
| Observational data | Non-processed data from 146 sites/grids from the WDCGG |
| Selection by standard scales | Data on the WMO standard gas scale (Zhao and Tans, 2006) were selected |
| Other selections | Observational data with small differences from model estimates were selected. |
| Uncertainty of observed data | Calculated from differences from smoothed data. Large uncertainties are given to missing or rejected data. |
| Transport Model | Offline transport model (JMA-CDTM, Sasaki et al., 2003) |
| Horizontal resolution | $2.5^\circ \times 2.5^\circ$ (global) |
| Vertical resolution | 32 layers (surface to 10 hPa) |
| Meteorology | JRA-25 (1985–2004, Onogi et al., 2007) and JCDAS (2005–present) |
| Transport scheme | Semi-Lagrangian (horizontal) and box scheme (vertical) |
| Vertical diffusion | Turbulent (Mellor and Yamada 1974), cumulus convection (Kuo, 1974) and shallow convection. |

data set enabled us to analyse $CO_2$ concentrations from 1985 to 2007. The main features of the analysis method are described in Table 2.

## 3. Experiments and validation

### *3.1. Estimated regional fluxes*

In this section, we show estimated fluxes in three experiments: CNTL, OBS and GV cases. In the CNTL case consisting of all processes of data selection, examined the contribution of meteorological variability. In the OBS case, Step 7 in Section 2.1.2 was skipped, so that only extreme outliers were excluded from the observational data. In the GV case, using processed GLOBALVIEW data, we did not apply the data selection method described earlier because the data were already quality controlled. Figure 4, which relates the global land and ocean fluxes to the ENSO index, defined as the sea surface temperature (SST) deviations from the climatological averages over a sliding 30-year period in the east tropical Pacific. The important feature is that there was no large difference in the results between the three cases. This shows that our data and site screening approach is replaceable to other data quality control systems in previous studies (e.g. GLOBALVIEW). After 1994, small (0.1–0.3 PgC $yr^{-1}$) differences are seen between CNTL case and GV, OBS cases. This may be due to the effect of data (not site) selection. The results also shows that the total land $CO_2$ fluxes increased following El-Niño events (1987–1988, 1997–1998 and 2002–2003), with delays of several months. This relationship was not observed in 1991–1992, due to the eruption of Mt. Pinatubo in June 1991 and the introduction of large amounts of aerosols into the atmosphere, thus affecting the climate. This study also shows that ocean global fluxes increased during El-Niño events, but that the peaks in 1997–1998

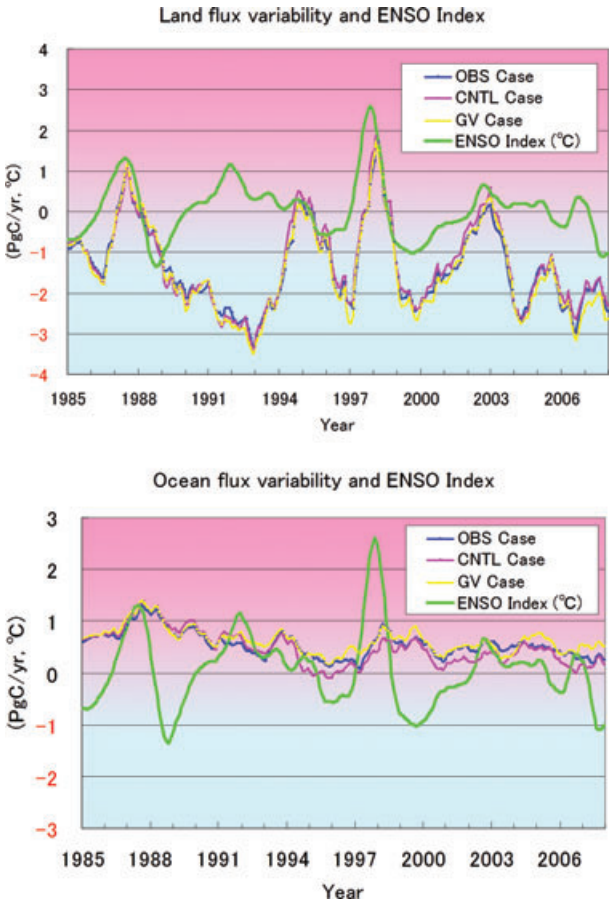


*Fig. 4.* Relationship between the variability in global land (upper) and ocean (lower) fluxes and the ENSO index [sea surface temperature (SST) deviations from the climatological mean in the east tropical Pacific].

in land flux were smaller than previously reported (Baker et al., 2006; Gurney et al., 2008), thus contributing to the increase in global $CO_2$ concentration. This may be due to the specifications of relatively strong vertical mixing in the lower troposphere in JMA-CDTM.

To evaluate the sensitivity of data selection in the flux estimation in a regional scale, the three methods (CNTL, OBS and GV cases) were compared. With respect to carbon fluxes, the global difference between the three cases was not very large, but there were some regional differences (Figs 5 and 6). In land regions, large differences were seen in less constrained regions where many data were rejected, for example Temperate Asia, Europe, Boreal Eurasia, Temperate North America and Tropical Africa. The interesting result was that the fluxes in Temperate Asia increased only in the CNTL case. The OBS and GV cases did not represent the features of observational data (summer minimum and winter maximum). In ocean regions, large differences were seen in the Northern Ocean, north Pacific, west tropical Pacific and south Indian Ocean. The large difference in the north Pacific, a well-constrained region, may be due to smaller rates of data selection at many sites/grids in this region, as shown in Fig. 2. Flux variability (standard deviation from the average) was generally the smallest in the OBS case and the largest in the CNTL case. The result is due to the features of observational data in each case. In the CNTL case, we used all observational data from the selected sites. In the GV case, we used all processed observational data from GLOBALVIEW. In the OBS case, we used only selected data from selected sites and the data rejection rate was relatively high in summer and winter. The estimated uncertainties of fluxes are summarized in Tables 3 and 4. The former indicates that the uncertainties for posterior sources did not differ much between the cases in many regions. The differences are only seen in the regions (Europe, Temperate Asia and Temperate North America) where many data are rejected in the OBS case. To validate such results, we needed to validate our results with independent observational data (e.g. continuous aircraft observation; Machida et al., 2008).

Carbon fluxes in the 1980s (1985–1990), 1990s (1991–2000) and 2000s (2001–2007) were estimated for different regions and for the whole globe. In the 1980s, the averaged land and ocean uptake without fossil fuel emission was 0.9 and 1.3 PgC $yr^{-1}$, respectively. The tropical emission was 1.1 PgC $yr^{-1}$, the northern uptake was 3.2 PgC $yr^{-1}$ and the southern flux was nearly zero. In the 1990s, due to the effect of the eruption of Mt. Pinatubo in June 1991, the total pre-subtracted flux was about 1.0 PgC $yr^{-1}$ larger than the increased amount of atmospheric $CO_2$. The northern and southern land regions contributed to the reduced fluxes. In the 2000s, as the total pre-subtracted flux increased, land and ocean uptake remained at the level of the 1990s. One important feature of this analysis is that the uptake by the Southern Ocean was weaker than the prior flux, as suggested previously (Gurney et al., 2002; Baker et al., 2006).

### 3.2. *Sensitivity tests*

As described in the previous sections, the analysis in this study features the use of non-processed (statistical monthly mean) observational data, including ship measurements and re-analysed meteorological data for a period longer than two decades. To evaluate the impacts of these features (data selection method, meteorological variability and ship observational data) and analysis precision, we assessed sensitivity by the following tests.

The first test, referred to as the 'CNTL' case, consisted of all processes of data selection and examined the contribution of meteorological variability. The second 'CLMT' case used the averaged results from the transport model for the whole analysis period as the transport coefficients. In the other tests, where the results of the transport model for specific years were used as a transport matrix, are referred to as the 'Y1997', 'Y1998', 'Y1999', 'Y2000' and 'Y2001' cases, based on the starting year of the transport model simulation with re-analysed meteorology.

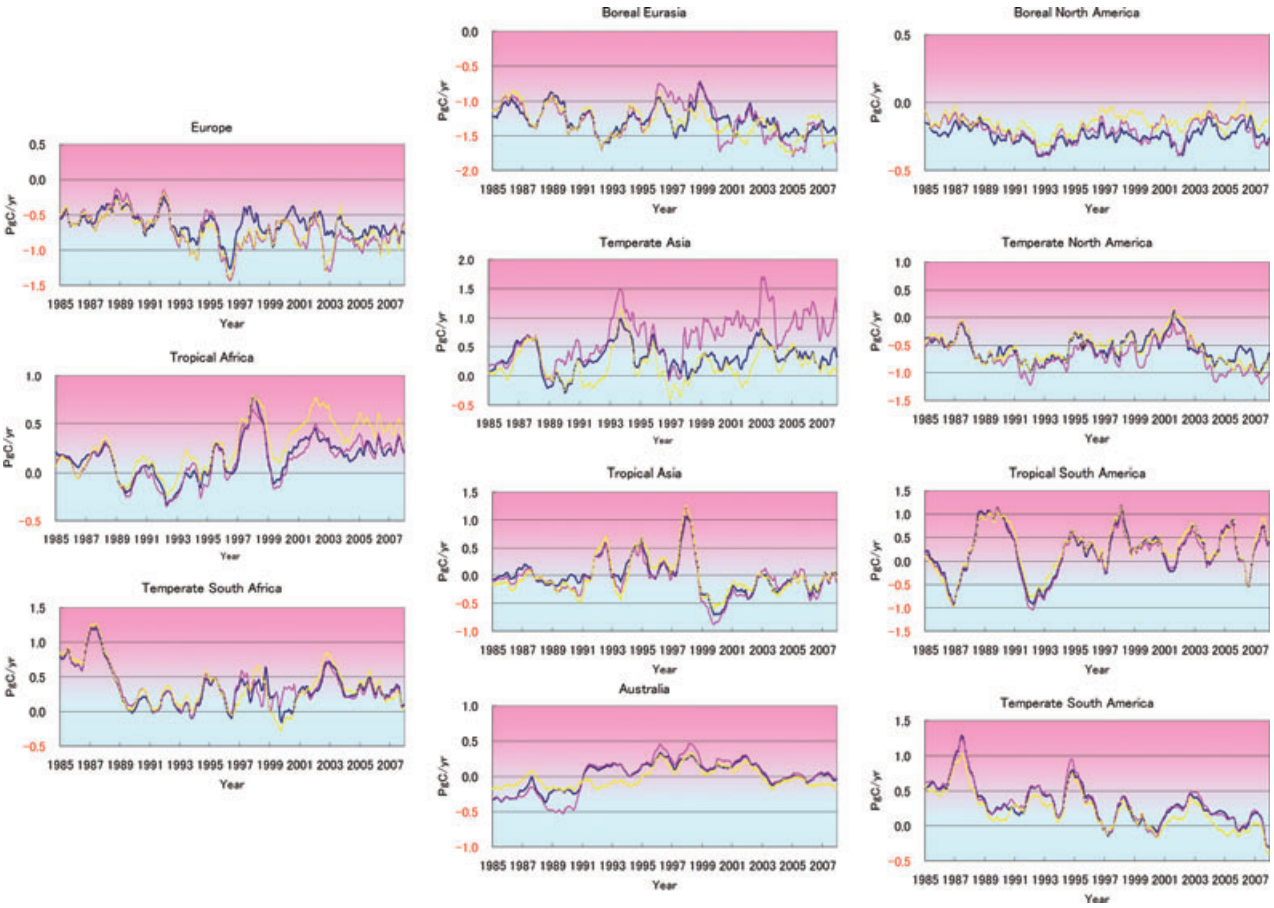


*Fig. 5.* Estimated $CO_2$ flux variabilities in land regions. The blue, red and yellow lines indicate the OBS, CNTL and GV cases, respectively.

The mismatches in $CO_2$ concentration between the observational data and the results of these tests are shown for different regions of the world in Table 5. The mismatch was the smallest in 'CLMT'. The mismatch in 'CNTL' was a little larger than that of 'CLMT', but smaller than the cases from 'Y1997' to 'Y2001'. This series of tests showed that model results are suboptimal when using specific years as a transport matrix. The reason why 'CLMT' provided similar mismatch to 'CNTL' may be due to some imperfection in the transport model, especially in vertical transport.

In the second set of calculations, several sites including 'Ryori (142°E, 39°N)', 'Minamitorishima (154°E, 24°N)', 'Yonagunijima (123°E, 24°N)', 'Iceland (20°W, 63°N)' and 'Easter Island (109°W, 27°S)' were removed from the 'CNTL' case to check the stability of the analysis and the contribution of observational sites. These stations are located in the area where the estimated fluxes are largely different between the CNTL and OBS cases described in the previous section.

Table 6 shows that data mismatches between the observation and analysis were mostly in ranges of 0.5–1.0 ppm. The mismatches were small in the oceans and Southern Hemisphere, but large in the land areas and Northern Hemisphere due to the error in pre-subtracted fluxes (especially biospheric flux). The mismatches became larger in all cases when specific sites were removed. The increase in data mismatch was significant when there were no other sites around the removed site as was the case in Iceland, whereas the mismatches did not change when many other sites existed around the removed site (e.g. Minamitorishima and Yonagunijima). The data mismatches were larger for land sites like Ryori, where analysed concentrations were sensitive to regional land flux, than for oceanic sites. The reason why the mismatch increased small in Easter Island may be that concentration variability is small in the Southern Hemisphere.

In the third set of calculations, 'gridded' ship sites were excluded from the 'CNTL' case. Posterior flux uncertainties were compared because no significant changes were expected in the distribution of $CO_2$ concentrations and data mismatches between the two cases. As are shown in Tables 3 and 7, posterior flux uncertainties became smaller than in the 'CNTL' case for all regions except Tropical Indian Ocean (+0.1%). The flux uncertainties were largely reduced in neighbouring (Boreal North America, Boreal Eurasia, Tropical Asia and Temperate Asia) and less constrained (Temperate South America, Boreal Eurasia and Southern Ocean) regions.

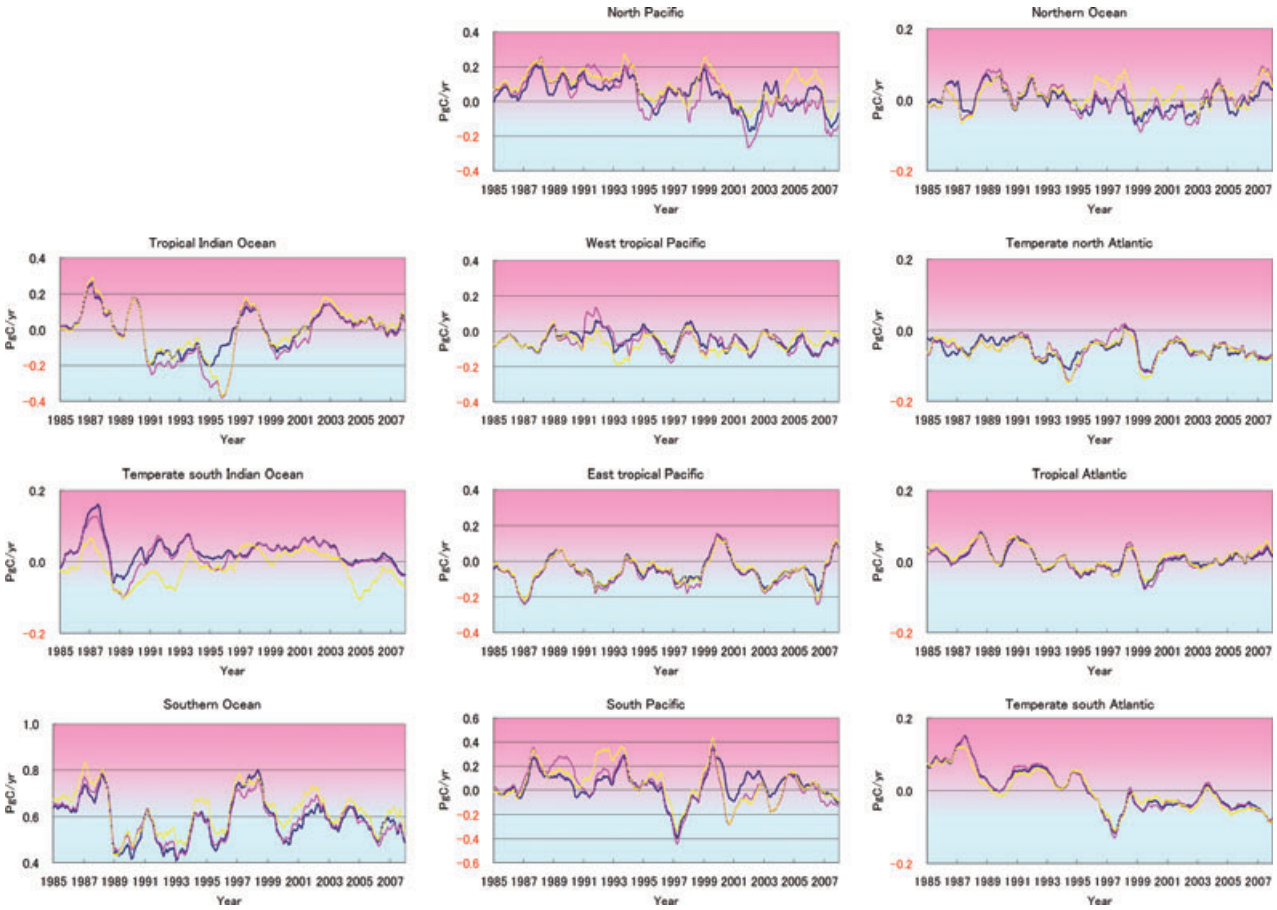


*Fig. 6.* Estimated $CO_2$ flux variabilities in ocean regions. The blue, red and yellow lines indicate the OBS, CNTL and GV cases, respectively.

### *3.3. Comparison with a similar study*

The monthly $CO_2$ concentrations in the 'CNTL' case were compared with those in CarbonTracker (Peters et al., 2007), which estimated carbon fluxes and concentrations at a high resolution from 2001 to 2007 using the Ensemble Kalman Filter (EnKF). In Fig. 7, the data mismatches in the 'CNTL' case were compared with those in CarbonTracker for different observational sites. The averaged data mismatches were similar to each other, 0.9 ppm for CNTL and 1.1 ppm for CarbonTracker and their geographical patterns did not differ greatly, except for northern land regions where data mismatches were smaller in CarbonTracker. This may be due to the higher spatial and temporal resolutions in carbon fluxes for CarbonTracker, especially in land regions.

## 4. Result and conclusion

We have shown here the ability to estimate global distributions of $CO_2$ concentrations over more than two decades with a new observational data and site screening technique. The estimated global $CO_2$ fluxes were quite similar to those shown previously (Stephens et al., 2007). Due to the specifications of our transport model regarding strong vertical mixing, the estimated uptake was relatively weak in the northern lands but relatively strong in the tropics. In the Southern Hemisphere and oceanic areas, $CO_2$ concentrations were estimated to precisions of 0.5–1.0 ppm. Our analysis is characterized by dependence on the use of observational data and on its selection of sites, especially the use of more marine boundary and less land sites. Theoretically, more land sites are included by adopting an analysis scheme with higher horizontal and temporal resolutions. This method of data selection may easily be introduced to other inversion method studies without developing new methods for site and data selection or data quality control.

The first experiment showed that interannual variability in meteorology was important in estimating $CO_2$ fluxes; the averaged climatological meteorology produced the equivalent results, given the current specifications of our transport model. Modified transport processes, especially a vertical transport scheme, would improve a performance in using interannual meteorology.

The second experiment showed the stability of our analysis. In areas of densely distributed observational sites, the precision

*Table 3.* Prior flux and posterior flux uncertainties for different regions (PgC $yr^{-1}$)

| Region name | Prior flux (biosphere and ocean) | Prior flux uncertainty | Posterior flux uncertainty (CNTL case) | Posterior flux uncertainty (OBS case) | Posterior flux uncertainty (without ship) |
|---|---|---|---|---|---|
| Boreal N. America | 0.0 | 0.35 | 0.29 | 0.29 | 0.31 |
| Temp. N. America | 0.0 | 0.84 | 0.60 | 0.62 | 0.64 |
| Tropical S. America | 0.0 | 1.34 | 0.99 | 0.99 | 1.02 |
| Temp. S. America | 0.0 | 0.87 | 0.79 | 0.79 | 0.79 |
| Tropical Africa | 0.0 | 0.77 | 0.73 | 0.73 | 0.73 |
| Temp. S. Africa | 0.0 | 0.93 | 0.81 | 0.81 | 0.83 |
| Boreal Eurasia | 0.0 | 0.70 | 0.41 | 0.42 | 0.45 |
| Temperate Asia | 0.0 | 0.79 | 0.59 | 0.62 | 0.65 |
| Tropical Asia | 0.0 | 0.60 | 0.44 | 0.45 | 0.47 |
| Australia | 0.0 | 0.32 | 0.28 | 0.28 | 0.29 |
| Europe | 0.0 | 0.70 | 0.50 | 0.53 | 0.55 |
| North Pacific | −0.50 | 0.28 | 0.24 | 0.24 | 0.25 |
| West Tropical Pacific | 0.15 | 0.20 | 0.18 | 0.19 | 0.18 |
| East Tropical Pacific | 0.47 | 0.22 | 0.19 | 0.19 | 0.19 |
| South Pacific | −0.23 | 0.38 | 0.32 | 0.32 | 0.32 |
| Northern Ocean | −0.44 | 0.16 | 0.14 | 0.14 | 0.14 |
| North Atlantic | −0.29 | 0.18 | 0.17 | 0.17 | 0.17 |
| Tropical Atlantic | 0.13 | 0.18 | 0.17 | 0.17 | 0.17 |
| South Atlantic | −0.13 | 0.20 | 0.19 | 0.19 | 0.19 |
| Southern Ocean | −0.89 | 0.46 | 0.27 | 0.27 | 0.27 |
| Tropical Indian | 0.12 | 0.26 | 0.24 | 0.24 | 0.24 |
| South Indian | −0.55 | 0.21 | 0.20 | 0.20 | 0.21 |
| Global | −2.16 | 2.76 | 0.31 | 0.31 | 0.31 |

of the analysis was stable, independent of an increase or decrease in the number of observational sites. However, in areas of sparse distribution, analysed $CO_2$ concentrations were sensitive to changes in the observational network especially in the northern hemisphere. To perform highly precise and stable analyses, data from other types of observation should be introduced, including continuous measurements on board aircraft and observations from satellites. The former produces observational data with a similar precision to observations on the surface (Machida et al., 2008), and the latter covers almost the whole world.

The third experiment showed that our analysis improved as observational data increased, beginning in 1993, and that our analysis was highly precise and stable especially in the Pacific area. To make such evaluations for other areas, other types of observational data should be introduced, as described earlier.

When we compared our analysis with that of CarbonTracker in terms of data mismatches, we found that the performance of our analytical method was sufficiently high. Considering the current observational network, the quality of our analysis was constantly high for more than 20 years.

## 5. Summary

A new technique has been developed to treat observational data in analysing global $CO_2$ concentrations. This technique does not require any other method for the quality control of observational data. Using this technique, non-processed observational data were easily analysed to estimate $CO_2$ fluxes and distributions for more than two decades in accordance with an inversion setup without developing a quality control system for observational data. This is an important feature of our analysis that researchers can make use of this technique with their own inversion setup. With the high-resolution transport model and flux regions, we can use observational data that were not available in the current inversion. However, we may result in rejecting site/data that should be used in the inversion due to imperfect inversion setups (e.g. transport model and prior surface fluxes). We need to improve further our inversion setup to obtain robust carbon fluxes. The estimated $CO_2$ fluxes were similar in quantity to those previously reported. In the 1990s, $CO_2$ uptakes totalled about 1.4 PgC $yr^{-1}$ on land areas and 1.8 PgC $yr^{-1}$ in the oceans. The uncertainty of global $CO_2$ flux estimation was about 0.3 PgC $yr^{-1}$. Estimated $CO_2$ uptakes were relatively weak in the northern hemisphere and strong in the tropics. These results are almost consistent with previous studies. The analysed $CO_2$ concentrations had precisions of about 0.5–1.0 ppm in the marine boundary layer. The mismatches between the observed and estimated concentrations were generally larger in northern land areas than in other areas. Several modifications would reduce the mismatches in our analysis. The first is the inclusion of

*Table 4.* Long-term averaged $CO_2$ flux estimates from this study for the 1980s (1985–1990), 1990s (1991–2000) and 2000s (2001–2007) in the two different experiments (PgC $yr^{-1}$)

| Region name | 1980s CNTL | 1990s CNTL | 2000s CNTL | 1980s OBS | 1990s OBS | 2000s OBS |
|---|---|---|---|---|---|---|
| Boreal N. America | −0.17 | −0.24 | −0.20 | −0.21 | −0.27 | −0.23 |
| Temp. N. America | −0.54 | −0.73 | −0.80 | −0.51 | −0.59 | −0.60 |
| Tropical S. America | 0.27 | 0.09 | 0.27 | 0.29 | 0.14 | 0.29 |
| Temp. S. America | 0.56 | 0.29 | 0.19 | 0.55 | 0.27 | 0.19 |
| Tropical Africa | 0.07 | 0.06 | 0.28 | 0.09 | 0.09 | 0.27 |
| Temp. S. Africa | 0.63 | 0.22 | 0.36 | 0.58 | 0.19 | 0.39 |
| Boreal Eurasia | −1.17 | −1.22 | −1.46 | −1.16 | −1.25 | −1.38 |
| Temperate Asia | 0.32 | 0.65 | 0.96 | 0.21 | 0.34 | 0.37 |
| Tropical Asia | −0.11 | 0.05 | −0.15 | −0.03 | 0.11 | −0.16 |
| Australia | −0.36 | 0.21 | 0.05 | −0.25 | 0.16 | 0.04 |
| Europe | −0.45 | −0.78 | −0.87 | −0.48 | −0.67 | −0.70 |
| North Pacific | −0.39 | −0.43 | −0.58 | −0.40 | −0.43 | −0.53 |
| West Tropical Pacific | 0.09 | 0.11 | 0.07 | 0.10 | 0.11 | 0.08 |
| East Tropical Pacific | 0.41 | 0.41 | 0.38 | 0.41 | 0.43 | 0.40 |
| South Pacific | −0.10 | −0.22 | −0.26 | −0.17 | −0.20 | −0.18 |
| Northern Ocean | −0.42 | −0.45 | −0.43 | −0.42 | −0.45 | −0.44 |
| North Atlantic | −0.33 | −0.35 | −0.35 | −0.33 | −0.35 | −0.35 |
| Tropical Atlantic | 0.16 | 0.12 | 0.14 | 0.16 | 0.12 | 0.14 |
| South Atlantic | −0.06 | −0.14 | −0.16 | −0.07 | −0.14 | −0.16 |
| Southern Ocean | −0.27 | −0.32 | −0.31 | −0.29 | −0.32 | −0.31 |
| Tropical Indian | 0.19 | −0.01 | 0.15 | 0.19 | 0.04 | 0.17 |
| South Indian | −0.54 | −0.52 | −0.53 | −0.51 | −0.51 | −0.53 |
| Northern Land | −2.01 | −2.31 | −2.37 | −2.16 | −2.43 | −2.54 |
| Northern Ocean | −1.14 | −1.22 | −1.36 | −1.15 | −1.23 | −1.32 |
| Tropical Land | 0.23 | 0.21 | 0.40 | 0.35 | 0.34 | 0.40 |
| Tropical Ocean | 0.85 | 0.63 | 0.74 | 0.87 | 0.71 | 0.79 |
| Southern Land | 0.83 | 0.73 | 0.60 | 0.88 | 0.62 | 0.61 |
| Southern Ocean | −0.97 | −1.20 | −1.27 | −1.04 | −1.17 | −1.19 |
| Northern total | −3.16 | −3.53 | −3.73 | −3.31 | −3.66 | −3.86 |
| Tropical total | 1.08 | 0.84 | 1.15 | 1.22 | 1.05 | 1.19 |
| Southern total | −0.14 | −0.47 | −0.66 | −0.16 | −0.55 | −0.57 |
| Total Land | −0.96 | −1.38 | −1.36 | −0.93 | −1.47 | −1.52 |
| Total Ocean | −1.26 | −1.79 | −1.88 | −1.33 | −1.69 | −1.73 |
| Global | −2.22 | −3.17 | −3.24 | −2.25 | −3.16 | −3.25 |
| Pre-subtracted Fossil Fuel | 5.83 | 6.41 | 7.64 | 5.83 | 6.41 | 7.64 |
| Atmospheric Increase | 3.61 | 3.33 | 4.32 | 3.67 | 3.30 | 4.31 |

*Table 5.* Data mismatch (observation–analysis) under different meteorological conditions

| Meteorology | CNTL | CLMT | Y1997 | Y1998 | Y1999 | Y2000 | Y2001 |
|---|---|---|---|---|---|---|---|
| Data mismatch (ppm) | 0.92 | 0.92 | 0.95 | 0.97 | 0.97 | 0.97 | 0.96 |

*Table 6.* Data mismatches when specific sites were removed from the CNTL run

| Site name | Ryori | Minamitorishima | Yonagunijima | Iceland | Easter Island |
|---|---|---|---|---|---|
| Data mismatch (CNTL) | 1.24 | 0.36 | 0.73 | 0.66 | 0.55 |
| Data mismatch in cases of removed site | 1.37 | 0.47 | 0.80 | 0.92 | 0.65 |

*Table 7.* Reduction of posterior flux uncertainties when ship measurements were included

| Region name | Boreal N. America | Tropical Asia | Boreal Eurasia | Temperate Asia | West Tropical Pacific | Europe |
|---|---|---|---|---|---|---|
| Reduction rate of posterior flux uncertainties (%) | 5.81 | 4.84 | 4.84 | 4.71 | 3.73 | 3.70 |

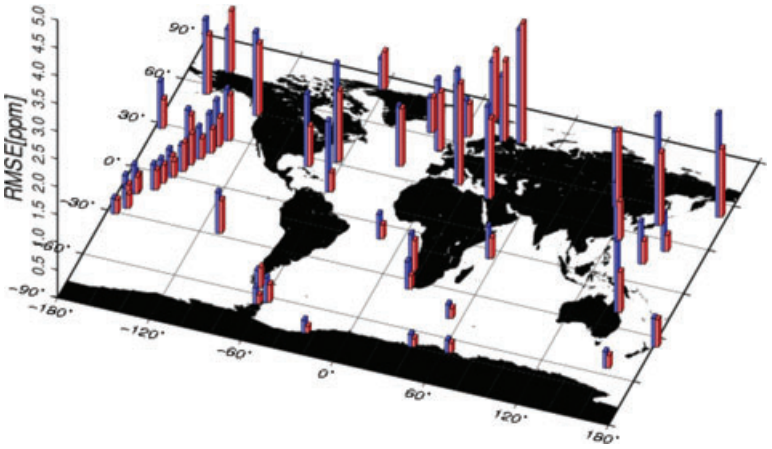


*Fig. 7.* Data mismatches in comparison with CarbonTracker (NOAA). The red and blue bars show the data mismatch of our analysis and CarbonTracker, respectively.

additional types of observational data by aircraft (Machida et al., 2008) and satellites (GOSAT). Secondly, inversion settings such as increased number of regions and reduced error in the pre-subtracted fluxes should improved the precision of the analysis. Thirdly, our transport model should interpolate the limited meteorological parameters to simulate the processes of tracer transport more precisely, by combining the current off-line model with GCM MJ98 (Shibata et al., 1999), which calculates GCM parameters without temporal or spatial interpolations. The EnKF would be an option for the analysis scheme to minimize computational costs and to realize a more precise analysis (Miyazaki, 2009).

## 6. Acknowledgments

We thank the NOAA/ESRL and other institutions (in Table 1) for making their observational data available. We thank R. Law, K. Gurney and P. J. Rayner for making the TDI model code available to us. We also thank two anonymous reviewers of this paper for their insightful and often challenging comments which helped improve the original version of this paper. The $CO_2$ data from Jubany is under an international agreement between the Italian PNRA and the Argentine DNA/IAA. This work was partly supported by grants from the National Science Foundation (OCE-9900310); the National Oceanic and Atmospheric Administration (NA67RJ0152, Amendment 30); the International Geosphere-Biosphere Programme/Global Analysis, Interpretation, and Modeling Project and the Global Carbon Project. This work was partly supported by the Environment Research and Technology Development Fund (A-0903) of the Ministry of the Environment, Japan.

## References

Apadula, F., Gotti, A., Pigini, A., Longhetto, A., Rocchetti, F. and co-authors. 2003. Localization of source and sink regions of carbon dioxide through the method of the synoptic air trajectory statistics. *Atmos. Environ.* **37**, 3757–3770.

Andres, R. J., Marland, G., Fung, I. and Matthews, E. 1996. A 1 × 1 degree distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. *Global Biogeochem. Cycles* **10**, 419–429.

Baker, D. F., Law, R. M., Gurney, K. R., Rayner, P., Peylin, P. and co-authors. 2006. TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional $CO_2$ fluxes, 1988–2003. *Global Biogeochemical Cycles* **20**, GB1002, doi:10.1029/2004GB002439.

Brenkert, A. L. 1998. Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis. (http://cdiac.esd.ornl.gov/ndps/ndp058a.html) Rep.NCP-058A, Carbon Dioxide Inf. Anal. Cent., doi:10.334/CDIAC/ffe.ndp058.2003.

Brunke, E.-G., Labuschagne, C., Parker, B., Scheel, H. E. and Whittlestone, S. 2004. Baseline air mass selection at Cape Point, South Africa: application of 222Rn and other filter criteria to $CO_2$. *Atmos. Environ.* **38**(33), 5693–5702.

Chamard, P., Thiery, F., di Sarra, A., Ciattaglia, L., De Silvestri, L. and co-authors. 2003. Interannual variability of atmospheric $CO_2$ in the Mediterranean: measurements at the island of Lampedusa. *Tellus* **55B**, 83–93.

Ciattaglia, L., Cundari, V. and Colombo, T. 1987. Further measurements of atmospheric carbon dioxide at Mt. Cimone, Italy: 1979–1985. *Tellus* **39B**, 13–20.

Ciattaglia, L., Colombo, T., Masarie, K. A. 1999. Continuous measurements of atmospheric $CO_2$ at Jubany Station, Antarctica. *Tellus* **51B**, 713–721.

Conway, T. J., Lang, P. M. and Masarie, K. A. 2007. Atmospheric carbon dioxide dry air mole fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968–2006, Version: 2007-09-19, ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/.

Eneroth K., Aalto, T., Hatakka, J., Holmen, K., Lourila, T. and co-authors. 2005. Atmospheric transport of carbon dioxide to a baseline monitoring station in northern Finland. *Tellus* **57B**, 366–374.

Engardt, M., Holmen, K. and Heintzenberg, J. 1996. Short-term variations in atmospheric $CO_2$ at Ny- Ålesund, Spitsbergen, during spring and summer. *Tellus* **48B**, 33–43.

Enting, I. 2002. *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, New York.

Francey, R. J., Steele, L.P., Langenfelds, R.L., Lucarelli, M. P., Allison, C. E. and co-authors. 1993. Global Atmospheric Sampling Laboratory (GASLAB): supporting and extending the Cape Grim trace gas programs. In: *Baseline Atmospheric Program (Australia) 1993*. Bureau of Meteorology and CSIRO Division of Atmospheric Research, Melbourne, Australia, 1996, 8–29.

Francey, R. J., Steele, L.P., Spencer, D.A., Langenfelds, R.L., Law, R. M. and co-authors. 2003. The CSIRO (Australia) measurement of greenhouse gases in the global atmosphere. In: *Report of the 11th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, Tokyo, Japan, September 2001* (eds. T. Sasaki and K. Suda). World Meteorological Organization Global Atmosphere Watch, 97–111.

Gurney, K., Law, R., Rayner, P. and Denning, A. S. 2000. *TransCom 3 Experimental Protocol*. Department of Atmospheric Science, Colorado State University, USA, 707.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D. and co-authors. 2002. Towards robust regional estimates of $CO_2$ sources and sinks using atmospheric transport models. *Nature* **415**(6872), 626–630.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D. and co-authors. 2004. Transcom 3 inversion intercomparison: model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochem. Cycles* **18**(1), GB1010, doi:10.1029/2003GB002111.

Gurney, K. R., Baker, D., Rayner, P. and Denning, S. 2008. Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated from atmospheric $CO_2$ inversions for the period 1980 to 2005, Global Biogeochem. *Cycles* **22**, GB3025, doi:10.1029/2007GB003082.

Haszpra, L., Barcza, Z., Bakwin, P. S., Berger, B. W., Davis, K. J. and co-authors. 2001. Measuring system for the long-term monitoring of biosphere/atmosphere exchange of carbon dioxide. *J. Geophys. Res.* **106D**, 3057–3070.

Higuchi, K., Worthy, D., Chan, D., Shashkov, A. 2003. Regional source/sink impact on the diurnal, seasonal and inter-annual variations in atmospheric $CO_2$ at a boreal forest site in Canada. *Tellus* **55B**, 115–125.

IPCC. 2007. Climate Change 2007. In: *The Physical Scientific Basis. Contribution of Working Group I to the Forth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Kuo, H. L. 1974. Further studies of the parameterization of the influence of cumulus convection on large scale flow. *J. Atmos. Sci.* **31**, 1232–1240.

Law R. M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I. and co-authors. 2008. Transcom Model simulation of hourly atmospheric $CO_2$: experimental overview and diurnal cycle results for 2002. Global Biogeochem. *Cycles* **22**, GB3009, doi:10.1029/2007GB003050.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Kondo, N. and co-authors. 2008. Worldwide measurements of atmospheric $CO_2$ and other trace gas species using commercial airlines. *Atmos. Oceanic. Technol.* **25**/**10**, 1744–1754.

Marland, G., Boden, T. A. and Andres, R. J. 2007. Global, regional, and national $CO_2$ emissions. In: *Trends: A Compendium of Data on Global Change*. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, TN, USA.

Matsueda, H., Inoue, H. Y. and Ishii, M. 2002. Aircraft observation of carbon dioxide at 8–13 km altitude over the western Pacific from 1993 to 1999. *Tellus* **54B**, 1–21.

Mellor, G. L. and Yamada, T. 1974. A hierarchy of turbulence closure models for planetary boundary layers. *J. Atmos. Sci.* **31**, 1791–1806.

Michalak, A. M., Hirsch, A., Bruhwiler, L., Gurney, K. R. , Peters, W. and co-authors. 2005. Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions. *J. Geophys. Res.* **110**, D24107, doi:10.1029/2005JD005970.

Miyazaki, K., 2009. Performance of a local ensemble transform Kalman filter for the analysis of atmospheric circulation and distribution of long-live tracers under idealized conditions. *J. Geophys. Res.* **114**, D19304, doi:10.1029/2009JD011892.

Nakazawa, T., Miyashita, K., Aoki, S. and Tanaka, M. 1991. Temporal and spatial variations of upper tropospheric and lower stratospheric carbon dioxide. *Tellus* **43B**, 106–117.

Navascues, B and Rus, C. 1991. Carbon dioxide observations at Izaña baseline station, Tenerife (Canary Islands): 1984–1988. *Tellus* **43B**, 118–125.

NOAA. 2007. GLOBALVIEW-CO2, Cooperative Atmospheric Data Integration Project – Carbon Dioxide, CD-ROM, NOAA CMDL, Boulder, Colorado.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S. and co-authors. 2007. The JRA-25 Reanalysis. *J. Meteor. Soc. Jpn.* **85**, 369–432.

Paramonova, N. N., Privalov, V. I. and Reshetnikov, A. I. 2001. Carbon dioxide and methane monitoring in Russia. Izvestia of Russian Academy of Science. *Atmos. Ocean Phys.* **37**(1), 38–43.

Patra, P. K., Law, R. M., Peters, W., Rödenbeck, C., Takigawa, M. and co-authors. 2008. TransCom model simulations of hourly atmospheric $CO_2$: analysis of synoptic scale variations for the period 2002–2003, Global Biogeochem. *Cycles* **22**, GB4013, doi:10.1029/2007GB003081.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J. and co-authors. 2007. An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *PNAS* **104**(48), 18925–18930.

Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y. and Field, C. B. 1997. The contribution of terrestrial sources and sinks to

trends in the seasonal cycle of atmospheric carbon dioxide. *Global Biogeochem. Cycles* **11**, 535–560.

Rayner, P. J., Enting, I. G., Francey, R. J. and Langenfelds, R. L. 1999. Reconstructing the recent carbon cycle from atmospheric $CO_2$, d13C and O2/N2 observations. *Tellus* **51B**, 213–232.

Rödenbeck, C., Houweling, S., Gloor, M. and Heimann, M. 2003. $CO_2$ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmos. Chem. Phys.* **3**, 2575–2659.

Rödenbeck, C., Conway, T. J. and Langenfelds, R. L. 2006. The effect of systematic measurement errors on atmospheric $CO_2$ inversions: a quantitative assessment. *Atmos. Chem. Phys.* **6**, 149–161.

Sasaki, T., Maki, T., Oohashi, S. and Akagi, K. 2003. Optimal sampling network and availability of data acquired at inland sites. *Global Atmos. Watch Rep. Ser.* **148**, 77–79.

Shibata, K., Yoshimura, H. H., Ohizumi, M., Hosaka, M. and Sugi, M. 1999. A simulation of troposphere, stratosphere and mesosphere with MRI/JMA98 GCM. *Papers Meteor. Geophys.* **50**, 15–53.

Sugi, M., Kuma, K., Tada, K., Tamiya, K., Hasegawa, N. and co-authors. 1990. Description and performance of the JMA operational global spectral model (JMA-GSM88). *Geophys. Magn.* **43**, 105–130.

Stephens, B., Gurney, K., Tans, P., Sweeney, C., Peters, W. and co-authors. 2007. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric $CO_2$. *Science* **316**, 1732–1735.

Tarantola, A. 1987. *Chapter 4 in: Inverse Problem Theory: Methods for Data Fitting and Parameter Estimation*. Elsevier, Amsterdam.

Takahashi, T., Wanninkhof, R. H., Feely, R. A., Weiss, R. F., Chipman, D. W. and co-authors. 1999. Net sea-air $CO_2$ flux over the global oceans: an improved estimate based on the sea-air $pCO_2$ difference. In: *Proceedings of the 2nd International Symposium: $CO_2$ in the Oceans, the 12th Global Environmental Tsukuba, 18–22 January 1999* (ed. Y. Nojiri). Tsukuba Center of Institutes, National Institute for Environmental Studies, Environment Agency of Japan, Tokyo.

Watanabe, F., Uchino, O., Joo, Y., Aono, M., Higashijima, K. and co-authors. 2000. Interannual variation of growth rate of atmospheric carbon dioxide concentration observed at the JMA's three monitoring stations: large increase in concentration of atmospheric carbon dioxide in 1998. *J. Meteorol. Soc. Jpn.* **78**, 673–682.

WMO. 2000. World Data Centre for Greenhouse Gases (WDCGG) Data Summary. WDCGG No. 22.

Zhao, C. L. and Tans, P. P. 2006. Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air. *J. Geophys. Res.* **111**, D08S09, doi:10.1029/2005JD006003.

## Supporting information

Additional supporting information may be found in the online version of this article:

**Appendix S1.** Location, data uncertainty and available data number of all sites/grids used for our analysis.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.