Evaluating and characterizing regional CO_2 fluxes estimated from satellite-based CO_2 data

Graduate School of Systems and Information Engineering
University of Tsukuba

March 2015

Hiroshi Takagi

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to my advisor, Dr. Tsuneo Matsunaga, for providing me with a great opportunity to pursue a doctorate degree. His careful, patient mentorship, detailed advice, and words of encouragement he gave me during the course of the program were inspiring. Thanks are also extended to my graduate committee members: Drs. Yoshiaki Osawa, Kunihiko Yoshino, Shun Watanabe, Akinobu Murakami, and Kenlo Nasahara. Their thoughtful suggestions, criticism, questions were all invaluable during the final stage of dissertation writing.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ACOS	The NASA Atmospheric CO_2 Observations from Space team
CASA	Carnegie Ames Stanford Approach
FLKS	Fixed-lag Kalman Smoother scheme
GFED	The Global Fire Emissions Database
GOSAT	Greenhouse Observing SATellite
GSNF	Growing season net fluxes
GV	GLOBALVIEW-CO2 data provided by NOAA
IPCC	The Intergovernmental Panel of Climate Change
JCDAS	JMA Climate Data Assimilation System
JMA	Japan Meteorological Agency
NEE	Net ecosystem exchange
NIES	National Institute of Environmental Studies
NIES-TM	NIES atmospheric tracer transport model
NOAA	The US National Oceanic and Atmospheric Administration
ODIAC	Open source Data Inventory of Anthropogenic CO_2 emission
OTTM	Ocean Tracer Transport Model
PPDF-S	Retrieval algorithm based on Photon Path length probability Density Function
RemoTeC	Retrieval algorithm developed by the Netherlands Institute for Space Research and Karlsruhe Institute of Technology, Germany
RMS	Root-mean-squared
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
SD	Standard deviation
SWIR	Short-wave infrared
TANSO-FTS	Thermal And Near infrared Sensor for carbon Observation Fourier transform spectrometer
TCCON	Total Carbon Column Observing Network
TransCom	The Atmospheric Tracer Transport Model Intercomparison Project
UoL-FP	University of Leicester full-physics retrieval algorithm
UR	Uncertainty reduction rate
VISIT	Vegetation Integrative SImulator for Trace gases

CHAPTER 1

Introduction – research background: understanding the global cycle of CO_2 using satellite remote sensing

The rapid atmospheric buildup of carbon dioxide (CO_2) observed over the past several decades [e.g. Keeling et al., 1976] raised a broad array of concerns about future climatic changes because of the role CO_2 plays in determining the Earth's heat budget [Ramanathan et al., 1987]. The Mauna Loa Observatory, operated by the US National Oceanic and Atmospheric Administration (NOAA), is one of the atmospheric observatories located around the globe for monitoring the long-term trend of atmospheric CO_2 levels. Figure 1.1 shows the result of the CO_2 measurement at the Observatory. Also shown in the figure are CO_2 records collected at NOAA's five other atmospheric monitoring stations. The figure indicates a steady rise in atmospheric CO_2 concentration around the globe, from the Antarctic to the Arctic. From the data collected at these global atmospheric monitoring stations between 2001 and 2010, the global-mean annual increase of CO_2 concentration was found to be 1.97 ppm [World Meteorological Organization, 2011]. (The unit *ppm* used here for the concentration of atmospheric CO_2 expresses how much volume of CO_2 in cm^3 occupies in 1 m^3 (1 million cm^3) of dry air (parts per million by volume).)

Based on an estimate for the total mass of the atmosphere (5.14×10^{18} kg

[Trenberth and Smith, 2004] or approximately 5,000 trillion tons), the global-mean annual CO_2 increase can be expressed in terms of the amount of CO_2 that was not absorbed and remains in the atmosphere. For the ten year period, the amount is calculated to be approximately 15.3 billion tons of CO_2 per year.

Since CO_2 in the atmosphere is inert, and the amount of CO_2 emitted through human activities, based on national fossil fuel consumption statistics, is known to be about 29.6 billion tons per year (estimate based on ODIAC anthropogenic emission inventory [Oda and Maksyutov, 2011]), the amount of CO_2 uptake by terrestrial vegetation and oceans can be estimated as about 14.3 billion tons per year. These figures point out that humans are emitting CO_2 approximately twice the amount terrestrial biosphere and ocean together are capable of absorbing in a year, thereby raising steadily the global atmospheric CO_2 concentrations.

As demonstrated above, it is possible to obtain an approximate global estimate of the amount of CO_2 exchanged between the atmosphere and the Earth's surface (denoted as surface CO_2 fluxes). However, with growing evidence of global climate change, as reported regularly by the Intergovernmental Panel of Climate Change (IPCC) [IPCC, 2013], there is an impending need, both scientifically and policy driven, to understand this global cycling of carbon in greater detail [Rayner and O'Brien, 2001]. Scientists and decision makers need to know the answers to overarching questions of 1) how anthropogenic CO_2 emissions are changing the global carbon cycle, 2) how policy and management decisions affect the level of atmospheric CO_2 concentration, and 3) how the rising atmospheric CO_2 levels, the associated changes in climate, and the carbon management decisions impact on ecosystems, biodiversity, and natural resources [Micharak et al., 2011]. Also, much research is needed as to the possibility that these human-induced changes in the global carbon cycle may eventually lead to shifting the Earth systems to new states, known as climate change tipping points, such as the ceasing of the global ocean conveyer belt and the melting of glaciers over Greenland. Gaining clear insight into these aspects is particularly important in projecting future changes in climate. Climate predictions rely upon estimates by multiple climate models that are forced with a common set of scenarios for atmospheric CO_2 levels [IPCC, 2013]. The development of reliable scenarios, essential for better future projections, is dependent on better answering the three questions listed above. Understanding the present and past state of the carbon cycle is the first yet critical step and lays a foundation for answering those intricate inquiries.

For this, there exist two approaches that give surface CO_2 flux estimates: the "bottom-up" and "top-down" approaches. CO_2 flux estimates by the bottom-up approach are obtained by summing up the estimates of CO_2 fluxes based on on-site observations, forestry statistics, fossil fuel consumption inventories, and land use change statistics, as well as those simulated by models of terrestrial biosphere and oceans. Although this method allows for the detailed estimation of CO_2 fluxes of particular regions, it may be difficult to obtain global scale estimates with it because detailed source data are available for particular parts of the globe. The top-down approach, on one hand, derives CO_2 fluxes from measured distributions of atmospheric CO_2 concentration, such as ones shown in Figure 1.1. This method is based on Bayesian inverse modeling, a statistical scheme used for inferring unknown values, such as locations on Earth, hypocenters of earthquakes, etc., from observations and a set of theoretical (or a priori) information on the value to be inferred (details on this approach is given in Chapter 2). This approach allows for global-scale CO_2 flux estimation, but there are issues associated with source data availability.

Attempts at studying the spatial distribution of CO_2 fluxes with the top-down approach have gathered pace in the late 1990s when individual estimates by different modeling systems were inter-compared in a series of research campaign called TransCom [e.g. Denning et al., 1999; Gurney et al., 2002]. In the third phase of the campaign, CO_2 flux estimates for 22 terrestrial and oceanic regions, based on data from 76 surface CO_2 monitoring sites, were compared against one another to gain insight into uncertainties inherent to the approach. Figure 1.2 shows the 22 global regions and the locations of the surface data providing sites used. The result showed that estimates for undersampled parts of the globe, particularly tropical latitudes, Africa, South America, and Asia (Figure 1.2), were associated with much larger uncertainties than those for temperate North America and Europe, where more data are available for the estimation [Gurney et al. 2002, 2004].

To augment the number and spatial coverage of the CO_2 data and reduce the flux uncertainties for the undersampled regions, it was suggested to use space-based spectral soundings of surface-reflected sunlight in the short-wave infrared (SWIR) wavelength range from which column-integrated CO_2 concentrations (X_{CO2}) can be retrieved [e.g.

Rayner and O'Brien, 2001; Houweling et al., 2004]. Rayner and O'Brien [2001] demonstrated that the satellite-based global X_{CO2} retrievals can reduce uncertainties in regional flux estimates substantially if data from the surface-based monitoring stations were augmented by the X_{CO2} retrievals with precisions of 1-2 ppm (~0.5%; on a regional scale with no zero systematic error, or "bias"). To this end, the Japanese Greenhouse Observing SATellite (GOSAT) was placed in orbit in early 2009. The satellite flies at an altitude of 666 km with a repeat cycle of 3 days. With an Earthward-looking Fourier transform spectrometer onboard, GOSAT takes global soundings of SWIR spectra in a raster scanning pattern (individual soundings are ~160 - 260 km apart in the cross-track direction), and approximately 60,000 X_{CO2} retrievals over clear-sky locations on land are obtained in a year.

With the advent of GOSAT, a new era has come to the estimation of surface CO_2 fluxes and the research of the global carbon cycle [e.g. Maksyutov et al., 2012; Chevallier et al., 2014; Basu et al., 2014]. As is always the case with any newly initiated data analyses, it is essential to evaluate and characterize first the new satellite-based CO_2 flux estimates and gain insight into the range of uncertainties associated with them before stepping into a stage in which the interpretation of those estimates is carried out. The objective of this study was therefore set as to evaluate the degree of contribution that GOSAT data make to the global surface CO_2 flux estimation and to elucidate sources of uncertainties associated with the flux estimates obtained and quantify them. In the chapters that follow, I will first give explanations on the top-down surface flux estimation scheme used for this study and its subsystems developed, as well as the first estimation results (Chapter 2), and present the utility of GOSAT data in the flux estimation (Chapter 3). In Chapters 4 and 5, I will present the results of investigating sources of uncertainties in the flux estimates using the developed system and the GOSAT data utility evaluation metric explained in Chapters 2 and 3; in Chapter 4, I will show how differences in X_{CO2} retrieval algorithms, as a source of the uncertainty, impacts the surface flux estimation, and then in Chapter 5 I will present how differences in X_{CO2} spatial coverage, another source of the uncertainty, affects the surface flux estimation. Finally in Chapter 6, I will sum up the new findings for gaining future research perspectives.

References

Basu, S., M. Krol, A. Butz, C. Clerbaux, Y. Sawa, T. Machida, H. Matsueda, C. Frankenberg, O. P. Hasekamp, and I. Aben (2014), The seasonal variation of the CO_2 flux over Tropical Asia estimated from GOSAT, CONTRAIL, and IASI, *Geophys. Res. Lett.*, **41**, doi:10.1002/2013GL059105.

Chevallier, F., P. I. Palmer, L. Feng, H., Boesch, C. W. O'Dell, and P. Bousquet (2014), Toward robust and consistent regional CO_2 flux estimates from in situ and spaceborne measurements of atmospheric CO_2, *Geophys. Res. Lett.*, **41**, doi:10.1002/2013GL058772.

Denning, S., et al. (1999), Three-dimensional transport and concentration of SF_6 - A model intercomparison study (TransCom 2), *Tellus*, **51B**, 266-297.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, **415**, 626–630.

Gurney, K. R., et al. (2004), TransCom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks. *Global Biogeochem. Cycles*, **18**, GB1010, doi:10.1029/2003GB002111.

Houweling, S., F.-M. Breon, I. Aben, C. Rödenbeck, M. Gloor, M. Heimann, and P. Ciais (2004), Inverse modeling of CO_2 sources and sinks using satellite data: A synthetic inter-comparison of measurement techniques and their performance as a function of space and time, *Atmos. Chem. Phys.*, **4**, 523–538, doi:10.5194/acp-4-523-2004.

IPCC (2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp, doi:10.1017/CBO9781107415324.

Keeling, C. D., et al. (1976), Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, *Tellus,* **28**(6), 538-551.

Maksyutov, S., et al. (2013), Regional CO_2 flux estimates for 2009–2010 based on GOSAT and ground-based CO_2 observations, *Atmos. Chem. Phys.,* **13,** 9351–9373, doi:10.5194/acp-13-9351-2013.

Micharak, A. M., et al. (2011), A US Carbon Cycle Science Plan, University Corporation for Atmospheric Research, Boulder, CO, USA.

Oda, T. and S. Maksyutov (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.

Ramanathan, V., et al. (1987), Climate-chemical interactions and effects of changing atmospheric trace gases, *Rev. Geophys.*, **25**, 1441-1482.

Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO_2 concentration data in surface source inversions. *Geophys. Res. Lett.*, **28**, 175−178.

Trenberth, K. E., and Smith, L. (2004), The mass of the atmosphere: a constraint on global analysis, *J. Climate*, **18**, 864-875.

World Meteorological Organization (2011), WMO Greenhouse Gas Bulletin, No.8, issued on November 19, 2012.

Figures

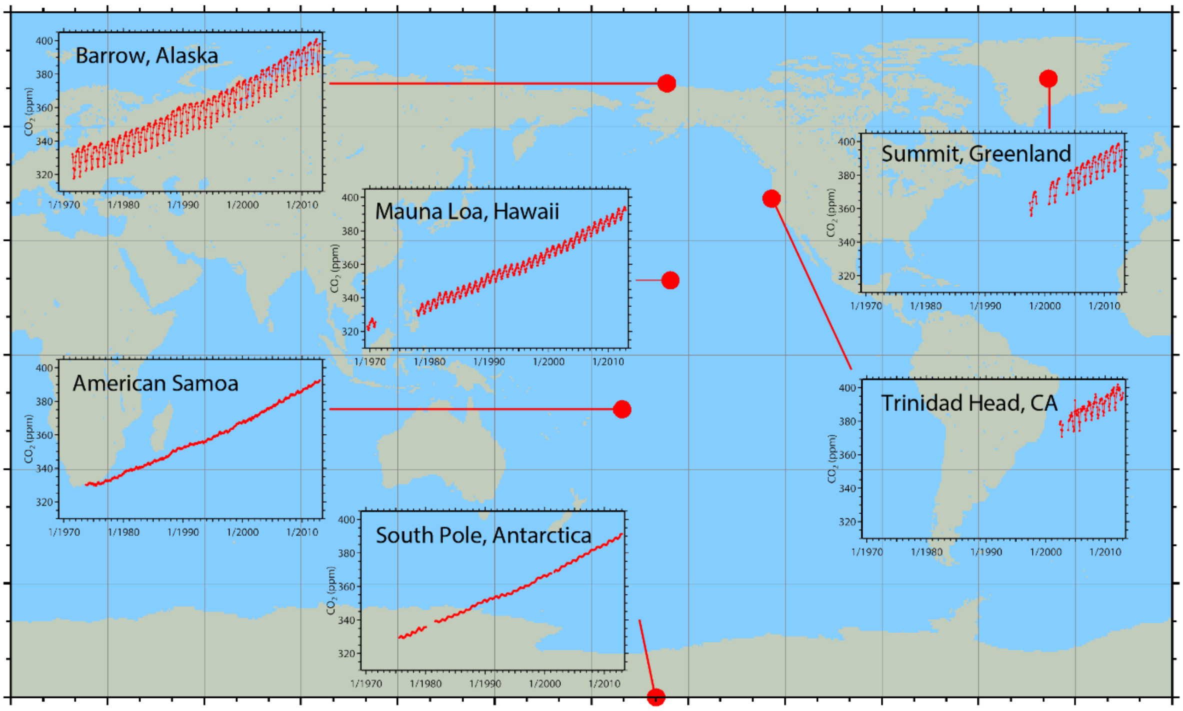

Figure 1.1. The time series of long-term CO_2 measurement taken at six atmospheric baseline observatories operated by the Earth System Research Laboratory of the US National Oceanic and Atmospheric Administration. The data were downloaded from: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/.

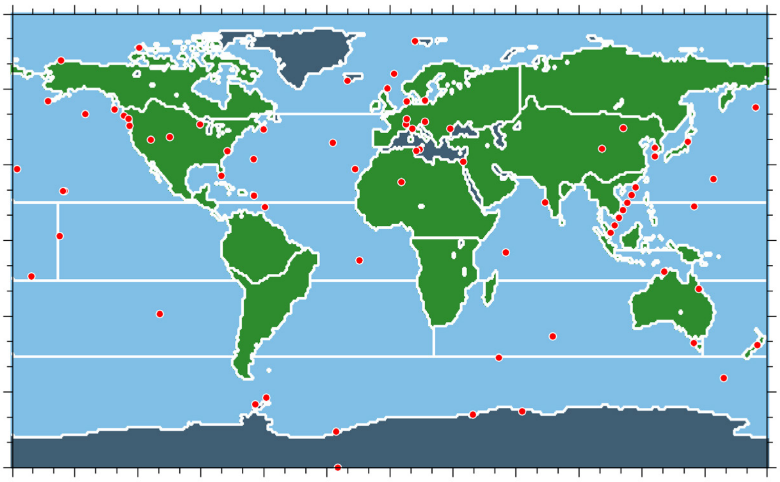

Figure 1.2. Boundaries of 22 terrestrial and oceanic regions used in the TransCom 3 flux intercomparison study. Red dots indicate the locations of 76 data providing sites used in the study.

CHAPTER 2

Estimation of regional CO_2 fluxes from GOSAT data

– approach and first result –

This study was made possible through collaborating with the following researchers:

Robert J. Andres[1], Dmitry Belikov[2,3,4], Isamu Morino[2], Tomohiro Oda[5,6], Makoto Saito[2], Ryu Saito[2,*], Osamu Uchino[2], Vinu Valsala[7], Tatsuya Yokota[2], Yukio Yoshida[2], and Shamil Maksyutov[2]

1 Oak Ridge National Laboratory, TN, USA

2 National Institute for Environmental Studies, Tsukuba, Japan

3 National Institute of Polar Research, Tokyo, Japan

4 Tomsk State University, Tomsk, Russian Federation

5 Colorado State University, CO, USA

6 Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, CO, USA

7 Indian Institute for Tropical Meteorology, Pune, India

* Now at Kokusai Kogyo Co. Ltd., Tokyo, Japan

2.1. Introduction

Here, explanations are given on the approach used throughout this study for estimating monthly regional CO_2 fluxes from both surface-based CO_2 data and GOSAT X_{CO2} retrievals. Descriptions are given on the following items involved in the regional CO_2 flux estimation: 1) inverse modeling scheme, 2) a priori flux data, 3) atmospheric tracer transport model, 4) GOSAT averaging kernel, 5) unit emission patterns, 6) flux and observation error covariance matrices, and 7) CO_2 concentration datasets. My involvement was in the development of the subsystems of the inverse modeling scheme in regard to items 4 through 6. Flux estimates obtained with the described approach are presented at the end of this chapter.

2.2. Descriptions of the flux estimation approach

2.2.1. Inverse modeling scheme

The top-down approach, or atmospheric inverse modeling, is a technique employed for inferring global surface CO_2 fluxes from the measurements of atmospheric CO_2 concentrations. The theoretical basis for the technique rests on Bayes' theorem [e.g. Tarantola, 2005], with which the "optimal" or a posteriori state of a set of parameters is deduced from a priori knowledge about those parameters and measured data values. In the case of estimating surface fluxes of CO_2, which is approximated to be chemically inert, the relationship between the measured data values and their theoretical predictions based on physical process modeling is linear. The relationship can be expressed in matrix form as

$$\boldsymbol{d}_{obs} = \mathbf{G}\boldsymbol{m} \tag{2-1}$$

where $\boldsymbol{d}_{obs}$ is the concentration vector recorded at measurement locations, and $\boldsymbol{m}$ denotes modeled CO_2 fluxes in predefined regions, respectively. **G** in Equation 2-1 represents the matrix of linear operators that maps the field of CO_2 fluxes onto that of concentrations. The elements of matrix **G** are given as changes in concentrations at each of measurement sites with respect to unit pulse emissions from each of the pre-defined regions. These elements, called the response functions, are obtained by running forward a set of unit pulse emissions (the basis functions) with an atmospheric tracer transport model [e.g. Rayner et al., 1999; Baker et al., 2006] (unit pulse emissions and atmospheric tracer transport model used in this study are explained in Sections 2.2.5 and 2.2.3). The magnitude of an element in the matrix, the "response" to a unit pulse emission, represents the degree of the contribution of individual observations to estimating a regional flux. The aim here is to find $\boldsymbol{m}$ that best describes $\boldsymbol{d}_{obs}$. Bayes' Theorem, formulated as

$$\mathrm{p}(\boldsymbol{m}|\boldsymbol{d}_{obs}) = \frac{\mathrm{p}(\boldsymbol{d}_{obs}|\boldsymbol{m})\ \mathrm{p}(\boldsymbol{m})}{\int \mathrm{p}(\boldsymbol{d}_{obs}|\boldsymbol{m})\ \mathrm{p}(\boldsymbol{m})\ \mathrm{d}\boldsymbol{m}}, \tag{2-2}$$

states that the a posteriori probability (probability of $\boldsymbol{m}$ given $\boldsymbol{d}_{obs}$, denoted as $\mathrm{p}(\boldsymbol{m}|\boldsymbol{d}_{obs})$), is equal to the probability of measurements (probability of $\boldsymbol{d}_{obs}$ given $\boldsymbol{m}$, $\mathrm{p}(\boldsymbol{d}_{obs}|\boldsymbol{m})$), times the a priori probability of $\boldsymbol{m}$ ($\mathrm{p}(\boldsymbol{m})$), normalized by the total probability. Assuming Gaussian error distributions, $\mathrm{p}(\boldsymbol{d}_{obs}|\boldsymbol{m})$ and $\mathrm{p}(\boldsymbol{m})$ are given as

$$\mathrm{p}(\boldsymbol{d}_{obs}|\boldsymbol{m}) = \frac{1}{\sqrt{2\pi\ \det\mathbf{C_D}}} \mathrm{e}^{-\frac{1}{2}(\mathbf{G}\boldsymbol{m}-\boldsymbol{d}_{obs})^{\mathrm{T}}\mathbf{C_D^{-1}}(\mathbf{G}\boldsymbol{m}-\boldsymbol{d}_{obs})} \quad \text{and} \tag{2-3}$$

$$\mathrm{p}(\boldsymbol{m}) = \frac{1}{\sqrt{2\pi\ \det\mathbf{C_M}}}\,\mathrm{e}^{-\frac{1}{2}(\boldsymbol{m}-\boldsymbol{m_p})^{\mathrm{T}}\mathbf{C_M^{-1}}(\boldsymbol{m}-\boldsymbol{m_p})}, \qquad (2\text{-}4)$$

respectively, where $\mathbf{G}\boldsymbol{m}$ denotes the expected values of $\boldsymbol{d_{obs}}$ (model prediction), and $\boldsymbol{m_p}$ is the a priori value of $\boldsymbol{m}$. $\mathbf{C_D}$ and $\mathbf{C_M}$ are the error covariance matrices of the observations and the a priori value, respectively (square matrices). Equations 2-3 and 2-4 together gives the posterior probability density as

$$\mathrm{p}(\boldsymbol{m}|\boldsymbol{d}_{obs}) \propto \mathrm{e}^{-\frac{1}{2}((\mathbf{G}\boldsymbol{m}-\boldsymbol{d}_{obs})^{\mathrm{T}}\mathbf{C_D^{-1}}(\mathbf{G}\boldsymbol{m}-\boldsymbol{d}_{obs})+\frac{1}{2}(\boldsymbol{m}-\boldsymbol{m_p})^{\mathrm{T}}\mathbf{C_M^{-1}}(\boldsymbol{m}-\boldsymbol{m_p}))}. \qquad (2\text{-}5)$$

The optimal state, ***m'***, is located at the center of this posterior probability density where the probability peaks out. ***m'*** can be found by minimizing the negative of the exponent in Equation 2-5 or the "cost function"

$$\mathbf{L}(m) = \frac{1}{2}(\mathbf{G}\boldsymbol{m} - \boldsymbol{d}_{obs})^{\mathrm{T}}\mathbf{C_D^{-1}}(\mathbf{G}\boldsymbol{m} - \boldsymbol{d}_{obs}) + \frac{1}{2}(\boldsymbol{m} - \boldsymbol{m_p})^{\mathrm{T}}\mathbf{C_M^{-1}}(\boldsymbol{m} - \boldsymbol{m_p}). \qquad (2\text{-}6)$$

Taking the derivative of **L** with respect to ***m*** gives

$$\frac{\partial \mathrm{L}(\boldsymbol{m})}{\partial \boldsymbol{m}} = \boldsymbol{m}(\mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\mathbf{G} + \mathbf{C_M^{-1}}) - \mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\boldsymbol{d}_{obs} + \mathbf{C_M^{-1}}\boldsymbol{m_p},$$

and, further, setting it to zero yields (the minimum of the cost function (2-6))

$$\begin{aligned}\boldsymbol{m}' &= (\mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\mathbf{G} + \mathbf{C_M^{-1}})^{-1}\left(\mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\boldsymbol{d}_{obs} + \mathbf{C_M^{-1}}\,\boldsymbol{m_p}\right), \text{ or} \\ &= \boldsymbol{m_p} + (\mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\mathbf{G} + \mathbf{C_M^{-1}})^{-1}\,\mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\left(\boldsymbol{d}_{obs} - \mathbf{G}\,\boldsymbol{m_p}\right). \qquad (2\text{-}7)\end{aligned}$$

Further, taking the derivative of **L** with respect to ***m*** for the second time gives

$$\frac{\partial^2 \mathrm{L}(\boldsymbol{m})}{\partial \boldsymbol{m}^2} = \mathbf{G}^{\mathrm{T}}\,\mathbf{C_D^{-1}}\,\mathbf{G} + \mathbf{C_M^{-1}}, \qquad (2\text{-}8)$$

which is the Hessian (the convexity) of the quadratic cost function (2-6).

As the cost function with respect to $\boldsymbol{m}$ is quadratic, the posterior probability density as presented in 2-5 is Gaussian, and can be expressed alternatively with the obtained optimal state $\boldsymbol{m'}$ at its center and the posterior covariance $\mathbf{C'_M}$ as

$$\mathrm{p}(\boldsymbol{m}|\boldsymbol{d}_{obs}) = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi}\,\mathrm{det}\mathbf{C'_M}}}\,\mathrm{e}^{-\frac{1}{2}\left(\boldsymbol{m}-\boldsymbol{m}'\right)^{\mathrm{T}}\mathbf{C'}_{\mathbf{M}}^{-1}(\boldsymbol{m}-\boldsymbol{m}')}, \text{ and}$$

$$\mathrm{p}(\boldsymbol{m}|\boldsymbol{d}_{obs}) \propto \mathrm{e}^{-\frac{1}{2}\left(\boldsymbol{m}-\boldsymbol{m}'\right)^{\mathrm{T}}\mathbf{C'}_{\mathbf{M}}^{-1}(\boldsymbol{m}-\boldsymbol{m}')}. \qquad (2\text{-}9)$$

The corresponding cost function is therefore written as

$$\mathbf{L}(\mathbf{m}) = \frac{1}{2}\left(\boldsymbol{m}-\boldsymbol{m}'\right)^{\mathrm{T}}\mathbf{C'}_{\mathbf{M}}^{-1}\left(\boldsymbol{m}-\boldsymbol{m}'\right). \qquad (2\text{-}10)$$

Taking the derivative of $\mathbf{L}$ twice with respect to $\boldsymbol{m}$ yields

$$\frac{\partial^2 \mathrm{L}(\boldsymbol{m})}{\partial \boldsymbol{m}^2} = \mathbf{C'}_{\mathbf{M}}^{-1}, \qquad (2\text{-}11)$$

which is the convexity of the quadratic cost function (2-10). With Equation 2-8, the posterior covariance matrix in Equation 2-11 (a square matrix) can be expressed as

$$\mathbf{C'_M} = \left(\frac{\partial^2 \mathrm{L}(\boldsymbol{m})}{\partial \boldsymbol{m}^2}\right)^{-1} = (\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})^{-1}. \qquad (2\text{-}12)$$

This equation can be rearranged as follows:

$$\begin{aligned}
\mathbf{C'_M} &= (\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})^{-1}(\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G}\,\mathbf{C_M} - \mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G}\,\mathbf{C_M} + \mathbf{I}) \\
&= (\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})^{-1}\Big((\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})\,\mathbf{C_M} - \mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G}\,\mathbf{C_M}\Big) \\
&= \mathbf{C_M} - (\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})^{-1}\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G}\,\mathbf{C_M} \\
&= \mathbf{C_M} - \mathbf{C}_{\mathrm{M}}\,\mathbf{G}^{\mathrm{t}}(\mathbf{G}\,\mathbf{C}_{\mathrm{M}}\,\mathbf{G}^{\mathrm{t}} + \mathbf{C}_{\mathrm{D}})^{-1}\mathbf{G}\,\mathbf{C_M}, \qquad (2\text{-}13)
\end{aligned}$$

since $\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,(\mathbf{G}^{\mathrm{T}}\,\mathbf{C}_{\mathbf{D}}^{-1}\,\mathbf{G} + \mathbf{C}_{\mathbf{M}}^{-1})^{-1} = \mathbf{C}_{\mathrm{M}}\,\mathbf{G}^{\mathrm{t}}(\mathbf{G}\,\mathbf{C}_{\mathrm{M}}\,\mathbf{G}^{\mathrm{t}} + \mathbf{C}_{\mathrm{D}})^{-1}$.

The right-hand side of Equation 2-12 shows how the observed data decrease the posterior error covariances.

The size of $\boldsymbol{m}$ in the present study was set to the number of flux estimation regions (64 regions) times the number of analyzed months. The 64 regions used in this study consist of 42 subcontinental-scale terrestrial regions and 22 ocean basins [Patra et al., 2005], which were defined by subdividing the original 22 land-ocean regions used in the TransCom 3 studies (Figure 1.2). The boundaries of these source regions are shown in Figure 2.1. The regions shaded with dark blue in the figure are not considered in the flux estimation. The dimension of matrix **G** is then determined as the size of $\boldsymbol{m}$ multiplied by that of vector $\boldsymbol{d}_{obs}$. For implementing matrix operations involved in Equation 2-7 efficiently, a variant of the fixed-lag Kalman Smoother scheme (FLKS), formulated by Bruhwiler et al. [2005], was employed. The basis for this scheme is the fact that in atmospheric tracer transport simulations, the signals of unit pulse emissions detected at measurement sites decay rapidly within the first few months and are blended into the background state thereafter. The idea is to obtain the a posteriori fluxes via estimating $\boldsymbol{m}$' incrementally with a subset of **G** and $\boldsymbol{d}_{obs}$ in a specified time-window. Using the FLKS setup with the same 64 region boundaries, Koyama et al. [2009] evaluated the influence that differences in the length of the time window have on a posteriori monthly flux estimates. Comparing results obtained using window lengths of 1 to 6 months, they concluded that a posteriori fluxes and their uncertainties estimated with three-month or longer windows converged quite strongly; Bruhwiler et al. [2005] arrived at a similar conclusion. Based on these findings, a window size of three month was chosen.

2.2.2 A priori fluxes

The a priori flux values stored in $\boldsymbol{m_p}$ (whose size is the same as that of $\boldsymbol{m}$) are comprised of four components: daily net ecosystem exchange (NEE) predicted by a terrestrial biosphere process model VISIT (Vegetation Integrative SImulator for Trace gases) [Ito, 2010; Saito M. et al., 2011]; monthly ocean-atmosphere CO_2 fluxes generated with an ocean pCO_2 data assimilation system run with the Ocean Tracer Transport Model (OTTM) [Valsala and Maksyutov, 2010]; monthly CO_2 emissions due to biomass burning stored in GFED (the Global Fire Emissions Database) version 3.1 [van der Werf et al., 2010]; and monthly anthropogenic CO_2 emissions obtained via merging the ODIAC (Open source Data Inventory of Anthropogenic CO_2 emission) high-resolution dataset [Oda and Maksyutov, 2011] and the Carbon Dioxide Information Analysis Center's monthly $1° \times 1°$ resolution dataset [Andres et al., 2011]. The spatial and temporal resolutions of these datasets are as follows: VISIT-predicted NEE: $0.5° \times 0.5°$ / daily; OTTM-based ocean flux: $1° \times 1°$ / monthly; GFED biomass burning emissions: $0.5° \times 0.5°$ / monthly; ODIAC anthropogenic emissions: $1° \times 1°$ (finer resolution data available) / monthly. Prior to the use in the forward concentration simulations, VISIT and GFED datasets were re-gridded to $1° \times 1°$. The estimation of NEE by VISIT is based on the Japan Meteorological Agency (JMA)'s JCDAS (JMA Climate Data Assimilation System) meteorological analysis data [Onogi et al., 2007].

2.2.3. Atmospheric tracer transport model

In this study, atmospheric tracer transport simulation necessary for constructing elements of matrix **G** and predicting concentrations at measurement locations was performed with version 08.1 of the National Institute for Environmental Studies (NIES) atmospheric tracer transport model (NIES-TM) [Belikov et al., 2011]. The tracer transport in NIES-TM is driven by JCDAS wind analysis data. The wind data are 6-hourly and are given on Gaussian horizontal grid T106 (320×160). Data for the height of the planetary boundary layer were taken from the interim reanalysis data provided by the European Center for Mid-range Weather Forecasts [Simmons et al., 2007]. Concentration simulation by NIES-TM is performed on a 2.5°×2.5° horizontal grid at 32 vertical levels between the surface and the top of the atmosphere (3 hPa). Validation against measurement made at twelve sites of the monitoring site of the Total Carbon Column Observing Network (TCCON) [Wunch et al., 2011a], where upward-looking high-resolution Fourier transform spectrometers are installed, showed that uncertainty associated with NIES-TM-simulated X_{CO2} is 0.2% of the concentration (~1 ppm) [Belikov et al., 2013].

2.2.4. Treatment of GOSAT averaging kernel in NIES-TM

To account for the vertical sensitivity of the GOSAT measurement in the prediction of GOSAT-based column-averaged concentrations, the averaging kernel, derived in the retrieval of X_{CO2}, was applied to each of the vertical concentration profiles simulated with NIES-TM. As described by Connor et al. [2008], a model-simulated X_{CO2} concentration $\boldsymbol{X^{m}_{CO_2}}$, which reflects the measurement vertical sensitivity, is given as

$$\boldsymbol{X^{m}_{CO_2}} = \boldsymbol{X^{a}_{CO_2}} \sum_{\mathbf{i}} (\boldsymbol{h}^{T}\mathbf{A})_{\mathbf{i}} \, (\boldsymbol{x_m} - \boldsymbol{x_a})_{\mathbf{i}}$$

where $\boldsymbol{X^{a}_{CO_2}}$ denotes a priori X_{CO2} values defined in the X_{CO2} retrieval, **A** is a matrix containing the CO_2 elements of the averaging kernel, $\boldsymbol{x_m}$ and $\boldsymbol{x_a}$ denote the elements of the modeled and a priori vertical CO_2 profiles, respectively. $\boldsymbol{h}$ is the pressure weighting function, a vector containing the dry air partial column abundance of each retrieval layer normalized to the total dry air column abundance. The calculation of the pressure weighting function was done as described in Appendix B of a report by Yoshida et al. [2009].

2.2.5. Unit emission patterns for constructing matrix G

For each of the monthly regional fluxes estimated, a concentration simulation was performed with NIES-TM in which a unit emission of 1 GtC region^{-1} yr^{-1} was released from that region for one month and transported forward until the end of the simulation period to sample responses at the location of every X_{CO2} retrieval. The spatial pattern of the 1 GtC region^{-1} yr^{-1} unit emission for each of the 42 land source regions (this is named the basis function), was defined as that of 31-yr-mean net primary productivity estimated by VISIT (1980-2010). Figure 2.2 shows the emission patterns for the 42 terrestrial regions. No spatial patterns were given to the unit emissions for the 22 ocean basins (spatially uniform). The sampled responses, named the response functions, were recorded in the columns of matrix **G**, which functions as a linear operator that relates concentrations with regional flux magnitudes.

2.2.6. Concentration datasets used for inverse modeling

The values assigned to the elements of vector $\boldsymbol{d}_{obs}$ are the surface-based GLOBALVIEW-CO2 (GV) data provided by NOAA [GLOBALVIEW-CO2] averaged monthly, and version 02.00 of GOSAT Level 2 X_{CO2} retrievals, distributed by the NIES GOSAT Project, that are gridded to 5°×5° cells and averaged monthly. Descriptions on these datasets are given below.

2.2.6.1. GLOBALVIEW data

The GV data are a product generated with a technique developed by Masarie and Tans [1995], which incorporates interpolated and/or extrapolated values with flask and in-situ continuous measurements such that the resulting smoothed concentration time series become seamless in time. A GV data file for a monitoring site contains 48 concentration values per year; for the estimation of monthly flux estimates in this study, these values were converted into monthly values. The reason behind the choice of GV data, instead of using simple averages of available flask and continuous observations in each month, as in a study by Rödenbeck et al. [2003], is to minimize the impact of temporal discontinuities that exist among those observations on the flux estimation.

Following the approach by Law et al. [2003], GV sites for the use in the flux estimation were selected by comparing GV data against concentrations predicted by NIES-TM over the one year analysis period. Sites whose root-mean-squared (RMS) model–observation misfits were less than 2 ppm were chosen. Altogether, 220 GV data time series were selected for this study (Table 2.1 shows the list of these sites). As an observation error estimate, the GV residual standard deviation (stored in the GV dataset) was assigned to each of the selected sites. Less weight was given at a GV site whose observational record completeness was less than 70% by tripling their data errors. Following Law et al. [2003], the minimum error for the GV data was set at 0.3 ppm.

2.2.6.2. GOSAT X_{CO2} retrievals

The TANSO (Thermal And Near infrared Sensor for carbon Observation) Fourier transform spectrometer (TANSO-FTS) is the main observational instrument aboard GOSAT, and measures surface-reflected sunlight and emitted thermal infrared radiation at wavelengths in the range 0.76–14.3 μm. The design and functions of the instrument are described in detail by Kuze et al. (2009). Sampled spectra recorded in the 0.76 μm oxygen absorption band and the 1.61 μm CO_2 absorption band were used in an earlier version of the NIES Level 2 operational retrieval algorithm (version 01; described by Yoshida et al., [2011]) to retrieve X_{CO2} global distributions. Those retrieved X_{CO2} values exhibited promising characteristics, including distinct north–south gradients and seasonal variability, but they were found to contain a significant negative bias of 8.85 ±4.75 ppm [Morino et al., 2011] when compared with reference data collected at the TCCON monitoring sites. Later, Uchino et al. [2012], using their lidar observations of aerosol particles, showed that assumptions made in version 01 of the retrieval algorithm on the vertical distributions of thin cirrus and aerosols are oversimplified, thereby contributing to the large bias. They proved that the issue could be mitigated significantly by the use of aerosol/cirrus optical properties retrieved simultaneously with spectra in the 2.06 μm band. Further, through investigating GOSAT spectra sampled over 2.5 yr, Yoshida et al. [2012] discovered a time-dependent degradation of TANSO-FTS's radiometric accuracy, which they successfully modeled for use in the retrieval algorithm implementation. These new findings, along with other improvements, were incorporated into the NIES Level 2 operational retrieval algorithm. The updated Level 2 X_{CO2} retrievals (version 02.00), processed from an improved GOSAT spectral dataset (Level 1B data, version 141.141, covering 14 months from June 2009 to July 2010) were shown to have a much smaller bias of -1.20 ± 1.97 ppm (the causes of the remaining bias, however, require further investigation).

Wunch et al. [2011b] made an attempt to assess and correct spatially- and temporally-varying biases in GOSAT X_{CO2} retrievals using an empirical regression model with which they correlated spurious variabilities in X_{CO2} retrievals with surface albedo, difference between the analyzed and retrieved surface pressure, airmass, and oxygen-band spectral radiance. A similar analysis is performed on the GOSAT Level 2 X_{CO2} retrievals [Inoue et al., in preparation], and the outcome of that effort will be reflected in the future updates of the X_{CO2} retrieval dataset. For this study, the bias was therefore corrected by raising each X_{CO2} value by the global mean GOSAT-TCCON difference of 1.20 ppm prior to the use in inverse modeling, assuming that the bias is uniform throughout the globe and the observation period.

Figure 2.3 shows the number of GOSAT X_{CO2} retrievals per each of 5°×5° cells counted during the months of August 2009, November 2009, February 2010, and May 2010. The distribution of the data number density changes with season owing to the occurrence of clear sky days and local solar zenith angle that determines the northern- and southern-most bounds of the GOSAT measurement. Note here that regions above 50° N latitude (the northern parts of North America and Eurasia) during fall and winter months saw very small numbers of GOSAT retrievals due to low local solar zenith angles therefore the flux inference for those regions during these months must rely on the GV data. Figure 2.4 displays GOSAT X_{CO2} retrievals in the form of input to the inverse modeling scheme (gridded to 5°×5° cells and averaged on a monthly time scale). Only the cells with three or more X_{CO2} retrievals per month are shown here. The monthly mean GV values are also shown in the figure in circles. The X_{CO2} bias correction was done prior to monthly averaging.

2.2.6.3. Model-simulated concentrations

The model-simulated concentration at each observation location of GV and GOSAT X_{CO2} was obtained by performing linear interpolation, in space and time, of the 2.5°×2.5° NIES-TM predicted concentration field (updated at a time step of 10-15 min in

NIES-TM). Monthly averaging of the predicted values was then followed. The monthly aggregation of individual predicted X_{CO2} values to a 5°×5° grid was done for grid cells that contain three or more X_{CO2} retrievals per month.

2.2.7. Prescribing error covariance matrices

The observation errors for the monthly mean X_{CO2} retrievals, specified in the diagonal elements of the error covariance matrix for the observations, $\mathbf{C_D}$, were determined as the standard deviations of GOSAT X_{CO2} retrievals found in each of the 5°×5° grid cells in a month. I took account of errors associated with the retrieval of X_{CO2} values and the forward atmospheric transport simulation by setting the minimum of the observation error for GOSAT X_{CO2} retrievals at 3 ppm, which consists of an uncertainty associated with the retrieval of GOSAT X_{CO2} (2 ppm) and that of the aforementioned forward X_{CO2} modeling (1 ppm). The $\mathbf{C_D}$ elements for GV data were set at the GV uncertainties described in Section 2.6.1.

The diagonal elements of the matrix $\mathbf{C_M}$ were prescribed as follows. The uncertainty of the terrestrial a priori flux was set at twice the standard deviation of the VISIT model monthly NEE (1°×1° resolution) values for the past 31 yr. The uncertainty of the oceanic a priori flux was determined as the RMS sum of the standard deviation of the OTTM-assimilated oceanic flux (1°×1° resolution) over a period between 2001 and 2009 and the mean square of differences between the OTTM-assimilated oceanic flux and climatological flux estimates by Takahashi et al. [2009].

In the TransCom 3 CO_2 inversion intercomparison, Gurney et al. [2003] assigned growing season net fluxes (GSNF; the sum of monthly-mean exchanges for months exhibiting net uptake) as terrestrial prior flux uncertainties (GSNF were based on NEE predicted by CASA (Carnegie Ames Stanford Approach) model [Randerson et al., 1997]). The reason behind it was that GSNF provide ecologically relevant upper bounds for annual-mean terrestrial flux. For oceanic fluxes, Gurney et al. [2003] set the uncertainties at 140% of the climatological net oceanic exchanges, which are approximately double the amount suggested by Takahashi et al. [2002]. The approach of using standard deviations of VISIT NEE and OTTM oceanic fluxes is similar to their case in finding reasonable upper limits of naturally varying fluxes and assigning them as boundaries in the flux estimation. These boundaries reflect natural variability in the past several decades (30 yr for terrestrial biosphere and 10 yr for ocean).

The off-diagonal elements of $\mathbf{C_D}$ and $\mathbf{C_M}$, i.e., the spatiotemporal covariances, were initially set at zero.

2.2.8. Flux estimation approach and its limitations

The above-described inverse modeling system gives the monthly estimates of surface CO_2 fluxes for the 42 sub-continental-scale terrestrial regions and 22 ocean basins of the globe, each of which is approximately 3000 km by 3000 km wide (Figure 2.1). The monthly regional CO_2 fluxes are derived by implementing matrix operations shown in Equation 2-7. As indicated in this equation, the regional flux estimates are obtained via "optimizing" or adjusting the a priori information on the monthly regional fluxes to be inferred (stored in $\boldsymbol{m_p}$); the term to the right of $\boldsymbol{m_p}$ in the right hand side of the equation corresponds to the adjustments made to the elements of $\boldsymbol{m_p}$ that are determined by the observed concentrations and response functions (Section 2.2.1) stored in $\boldsymbol{d_{obs}}$ and **G** (values are monthly-averaged in the modeling system), respectively, along with the magnitudes of covariances for the observations ($\mathbf{C_D}$) and a priori fluxes ($\mathbf{C_M}$). The response functions for individual observations are determined by atmospheric transport simulated with NIES-TM and the basis function (unit emission patterns) pre-defined regionally based on the VISIT-predicted strength of net primary productivity in each terrestrial regions (patterns for ocean basins are flat). The emissions due to fossil fuel and biomass burning, two of the four components of the a priori flux, are handled as given in the flux estimation. Thus, the adjustments to the a priori flux $\boldsymbol{m_p}$ are made with respect to the terrestrial biosphere and ocean fluxes.

The optimization in the inverse modeling before the advent of GOSAT, as in the TransCom 3 study campaign in the late 1990s, was performed on fluxes of regions that are much wider in area than those used in this study (22 global regions shown in Figure 1.2; approximately 7000 km by 7000 km wide). Figure 1.2 also shows how the 76 GV data providing stations used in the TransCom flux estimation are distributed among the 22 large regions; the horizontal distances among the 76 GV stations ranged from a few hundred kilometers (some stations in the US) to several thousand kilometers (stations over the under-sampled continents such as Africa and South America), indicating unevenness in the station distribution. Ideally, for the purpose of sampling CO_2 as spatially evenly as possible in the horizontal direction, it is desirable to locate the surface stations in a mesh. However, this is quite difficult because of challenges and issues in building, staffing, and maintaining financially new stations, particularly in the under-sampled regions of the globe. Since the frequency of the current CO_2 flask and in-situ sampling, which ranges about twice a month to several times in a second, is site-dependent, unevenness also exists in the temporal direction. GOSAT, launched to complement the surface-based measurements, does make spectral measurement in a mesh-like, raster-scanning pattern with a repeat cycle of three days. Historically, with the expectation that the overall CO_2 data volume would be significantly increased by the satellite, the 22 regions used in the TransCom study were further sub-divided into the 64 regions, as adopted by Maksyutov et al. [2013] and this study. Despite the overall data number leap, the horizontal distribution of the retrieved X_{CO2} can be space- and time-dependent thereby uneven because the spectral measurement by GOSAT can be perturbed by local clouds and aerosols, and the X_{CO2} retrieval is only possible for locations where the local solar zenith angle, which changes with season, is less than 70°.

To reduce the impact of the potential spatiotemporal unevenness in the CO_2 data distribution on the flux estimation, GOSAT X_{CO2} retrievals were gridded to 5°×5° cells and averaged on a monthly basis. This way, the horizontal data number distribution is regularized; the downside of it is that the X_{CO2} information content may be reduced via the data aggregation. The temporal consistency between the CO_2 data used and the fluxes estimated (both are monthly) is maintained in the present inverse modeling setup, but if the temporal resolution of the flux estimates were to be improved (e.g., to weekly or three-day estimates), then the limiting, bottleneck factor in resolving the regional fluxes would be the low temporal resolution of the surface CO_2 measurements.

Another factor in limiting the current flux estimation capability is the coarse-resolution modeling of atmospheric CO_2 transport. The prediction of measured and retrieved CO_2 concentrations in the present system with NIES-TM is performed on a 2.5° ×2.5° grid (a cell near the equator is approximately 250 km × 250 km wide); the grid size used is very close to what are adopted by the many other existing atmospheric transport models used for the flux estimation (lists of the existing transport models are found in the reports by Patra et al. [2011] and Niwa et al. [2011]). Concentration simulations on finer grids allow for accounting for smaller-scale details in the atmospheric transport and dispersion, but the increase in the computational cost is significant and cannot be ignored; Belikov et al. [2011] reported that a doubling of the horizontal resolution of NIES-TM (from 2.5°×2.5° to 1.25°×1.25°) leads to an increase in the computational time of about 5 times, and a quadrupling (0.625°×0.625°) requires 37 times more time than the current 2.5°×2.5° simulation. The forward concentration simulations required for the flux estimation over the one-year period lasted 4 days (single-core jobs run with Intel Xeon E5-4600 series processors on SGI UV20 servers installed at NIES); performing the same simulation on the doubled and quadrupled grids, based on the reported computing cost estimation, may require ~20 days and 148 days, respectively.

Balancing the computing cost and the efficiency in obtaining the end results is an issue here.

2.3. Flux estimation result

Using 14-month-long GOSAT Level 2 X_{CO2} retrievals (version 02.00) and the GV data in the 3-month-window FLKS scheme, monthly fluxes were inferred for the 64 subcontinental regions for 12 months between June 2009 and May 2010. A total of 9106 observations were available for the estimation, of which 6125 were gridded monthly-mean GOSAT X_{CO2} retrievals and 2981 were monthly-mean GV data. The monthly breakdown of the X_{CO2} number statistics are found in Table 2.2 (their spatial distributions are shown in Figure 2.7). Figure 2.5 presents the monthly posteriori fluxes for the months of August 2009, November 2009, February 2010, and May 2010 (results for the other months are found in Figure 2.6). Values in the unit of gC m^{-2} day^{-1} are shown. Positive and negative values, as color-coded in the figure, indicate whether a region served as a net source of CO_2 (net emission) or a sink of the gas (net absorption) for a given month. It can be seen in the figure that regions with net sink are predominant over the boreal regions of the North America and Eurasia in August 2009 (summer in the Northern Hemisphere) during which CO_2 uptake by forest via photo synthesis is maximum. The uptake then weakens during the fall and winter months of the Northern Hemisphere, and gradually comes back in the spring season (May 2010). The opposite is found in the high latitude bands of the Southern Hemisphere.

To indicate which regional fluxes are adjusted most by the surface and satellite CO_2 data in this one-year flux estimation, I present in Figure 2.8 the difference between the a posteriori fluxes (net) and the corresponding a priori values to which the optimization was performed. The values are shown as the a posteriori minus the a priori values in GtC region^{-1} year^{-1} (departure from the a priori value). It turned out that the optimization, on an annual time scale, lead to more CO_2 outgassing in most of the tropics (tropical America, tropical Africa, tropical Asia, and northern Australia), western United States (Regions 5 and 7), Eastern Eurasia (Regions 26 and 32), and middle South America (Regions 15 and 16). The optimization on one hand resulted in more CO_2 uptake in north-eastern US (Region 8), western Europe (Regions 39 and 41), northern Eurasia (Regions 25, 27, and 28), central Eurasia (Regions 30 and 31), and the high-latitudinal regions of the Southern Hemisphere (Regions 13, 14, 21, and 36). These terrestrial adjustments are in a range between -0.5 and 0.5 GtC region^{-1} year^{-1} (a 0.5 GtC region^{-1} year^{-1} emission is equivalent to about twice as much the GFED-estimated biomass-burning emissions from Region 17 (western tropical Africa) in a year).

References

Andres, R., Gregg, J. S., Losey, L., Marland, G., and Boden, T. (2011), Monthly, global emissions of carbon dioxide from fossil fuel consumption, *Tellus B*, **63**, 309–327.

Baker, D. F., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles,* **20***, GB1002, doi:10.1029/2004GB002439.*

Belikov, D. A., Maksyutov, S., Miyasaka, T., Saeki, T., Zhuravlev, R., and Kiryushov, B. (2011), Mass-conserving tracer transport modelling on a reduced latitude-longitude grid with NIES-TM, *Geosci. Model Dev.*, **4**, 207–222, doi:10.5194/gmd-4-207-2011.

Belikov, D. A., Maksyutov, S., Sherlock, V., Aoki, S., Deutscher, N. M., Dohe, S., Griffith, D., Kyro, E., Morino, I., Nakazawa, T., Notholt, J., Rettinger, M., Schneider, M., Sussmann, R., Toon, G. C., Wennberg, P. O., and Wunch, D. (2013), Simulations of column-averaged CO_2 and CH_4 using the NIES TM with a hybrid sigmaisentropic vertical coordinate, *Atmos. Chem. Phys.*, **13**, 1713–1732, doi:10.5194/acp-13-1713-2013.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P. (2005), An improved Kalman Smoother for atmospheric inversions, *Atmos. Chem. Phys.*, **5**, 2691–2702, doi:10.5194/acp-5-2691-2005.

Connor, B. J., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D. (2008), Orbiting Carbon Observatory: inverse method and prospective error analysis, *J. Geophys. Res.*,

113, D05305, doi:10.1029/2006JD008336.

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

Gurney, K. R., et al. (2003), Transcom 3 CO_2 Inversion Intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus* B, **55**, 555–579.

Inoue M. et al., Bias correction and evaluation of GOSAT SWIR X_{CO2} and X_{CH4} using TCCON data and aircraft measurements, in preparation

Ito, A. (2010), Changing ecophysiological processes and carbon budget in East Asian ecosystems under near-future changes in climate: Implications for long-term monitoring from a process-based model, *J. Plant Res.*, **123**, 577–588.

Koyama, Y., Maksyutov, S., and Valsala, V. (2009), Model analysis of observational data of the atmospheric tracers for the estimation of greenhouse gases in East Asia, Global Environment Research Account for National Institutes in FY2008, Ministry of the Environment, Japan, 49–95.

Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki (2009), Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Appl. Opt*. **48**, 6716-6733.

Law, R. M., et al. (2003), TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, *Tellus*, **55B**(2), 580– 595.

Masarie, K. A. and Tans, P. P. (1995), Extension and Integration of Atmospheric Carbon

Dioxide Data into a Globally Consistent Measurement Record, *J. Geophys. Res.*, **100**, 11593–11610.

Morino, I., et al. (2011), Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, *Atmos. Meas. Tech.*, **4**, 1061–1076, doi:10.5194/amt-4-1061-2011, 2011.

Niwa, Y., P. K. Patra, Y. Sawa1, T. Machida, H. Matsueda1, D. Belikov, T. Maki, M. Ikegami, R. Imasu, S. Maksyutov, T. Oda, M. Satoh, and M. Takigaw2. (2011), Three-dimensional variations of atmospheric CO_2: aircraft measurements and multi-transport model simulations, *Atmos. Chem. Phys.*, **11**, 13359–13375, doi:10.5194/acp-11-13359-2011.

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.

Onogi, K., et al. (2007), The JRA-25 reanalysis. *J. Meteor. Soc. Japan*, **85**, 369–432.

Patra, P. K., Ishizawa, M., Maksyutov, S., Nakazawa, T., and Inoue, G. (2005), Role of biomass burning and climate anomalies for land-atmosphere carbon fluxes based on inverse modeling of atmospheric CO_2, *Global Biogeochem. Cycles*, **19**, GB3005, doi:10.1029/2004GB002258.

Patra, P. K., et al. (2011), TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere, *Atmos. Chem. Phys.*, **11**, 12813–12837, doi:10.5194/acp-11-12813-2011.

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., Francey, R., and Langenfelds, R. (1999), Reconstructing the recent carbon cycle from atmospheric CO_2, ^{13}C and O_2/N_2 observations, *Tellus B*, **51**, 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**, 1919–1964, doi:10.5194/acp-3-1919-2003.

Saito, M., Ito, A., and Maksyutov, S. (2011), Evaluation of biases in JRA-25/JCDAS precipitation and their impact on the global terrestrial carbon balance, *J. Climate*, **24**, 4109–4125.

Simmons, A., Uppala, S., Dee, D., and Kobayashi, S. (2007), ERAInterim: New ECMWF reanalysis products from 1989 onwards, ECMWF, Newsletter No. 110 –Winter 2006/2007, 25–35.

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, N., Bates, N.,Wanninkhof, R., Feely, R. A., Sabine, C., Olafsson, J., and Nojiri, Y. (2002), Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep-Sea Res. II*, **49**, 1601–1622.

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

Tarantola, A. (2005), Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, Philadelphia, USA, 342 pp.

Uchino, O., Kikuchi, N., Sakai, T., Morino, I., Yoshida, Y., Nagai, T., Shimizu, A., Shibata, T., Yamazaki, A., Uchiyama, A., Kikuchi, N., Oshchepkov, S., Bril, A., and Yokota, T. (2012), Influence of aerosols and thin cirrus clouds on the GOSAT-observed CO_2: a case study over Tsukuba, *Atmos. Chem. Phys.*, **12**, 3393–3404, doi:10.5194/acp-12-3393-2012.

Valsala, V. and Maksyutov, S. (2010), Simulation and assimilation of global ocean pCO_2 and air-sea CO_2 fluxes using ship observations of surface ocean pCO_2 in a simplified biogeochemical offline model, *Tellus B*, **62**, 821–840.

van derWerf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T. (2010), Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), *Atmos. Chem. Phys.*, **10**, 11707–11735, doi:10.5194/acp-10-11707-2010.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R. A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock, V., and Wennberg, P. O. (2011a), The Total Carbon Column Observing Network: *Phil. Trans. R. Soc.* A, **369**, 2087–2112, doi:10.1098/rsta.2010.0240.

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyr¨o, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C. (2011b), A method for evaluating bias in global measurements of CO_2 total columns from space, *Atmos. Chem. Phys*., **11**, 12317–12337, doi:10.5194/acp-11-12317-2011.

Yoshida, Y., Oguma, H., Morino, I., Suto, H., Kuze, A., and Yokota, T. (2009) Mountaintop observation of CO_2 absorption spectra using a short wavelength infrared Fourier transform spectrometer, *Appl. Optics,* **49**, 71–79.

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T. (2011), Retrieval algorithm for CO_2 and CH_4 column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, *Atmos. Meas. Tech.,* **4**, 717–734, doi:10.5194/amt-4-717-2011.

Yoshida, Y., Kikuchi, N., and Yokota, T. (2012), On-orbit radiometric calibration of SWIR bands of TANSO-FTS onboard GOSAT, *Atmos. Meas. Tech.*, **5**, 2515–2523, doi:10.5194/amt-5-2515-2012.

Tables

Table 2.1. List of GLOBALVIEW sites used for this study (220)

SITE	LAT	LON	SITE	LAT	LON
ABP_01D0	-12.77	-38.17	NHA035_01P2	42.95	-70.63
AIA005_02D2	-40.53	144.30	NHA045_01P2	42.95	-70.63
AIA015_02D2	-40.53	144.30	NHA055_01P2	42.95	-70.63
AIA025_02D2	-40.53	144.30	OPW_01D0	48.25	-124.42
AIA035_02D2	-40.53	144.30	ORL015_11D2	47.80	2.50
AIA045_02D2	-40.53	144.30	ORL025_11D2	47.80	2.50
AIA055_02D2	-40.53	144.30	ORL035_11D2	47.80	2.50
AIA065_02D2	-40.53	144.30	PFA015_01P2	65.07	-147.29
ALT_01D0	82.45	-62.51	PFA025_01P2	65.07	-147.29
ALT_06C0	82.45	-62.51	PFA035_01P2	65.07	-147.29
AMS_01D0	-37.95	77.53	PFA045_01P2	65.07	-147.29
AMS_11C0	-37.95	77.53	PFA055_01P2	65.07	-147.29
AMT012_01C3	45.03	-68.68	PFA065_01P2	65.07	-147.29
AMT107_01C3	45.03	-68.68	PFA075_01P2	65.07	-147.29
ASC_01D0	-7.97	-14.40	POCS35_01D1	-35.00	180.00
ASK_01D0	23.18	5.42	POCS30_01D1	-30.00	-176.00
AVI_01D0	17.75	-64.75	POCS25_01D1	-25.00	-171.00
AZR_01D0	38.77	-27.38	POCS20_01D1	-20.00	-174.00
BHD_01D0	-41.41	174.87	POCS15_01D1	-15.00	-171.00
BHD_15C0	-41.41	174.87	POCS10_01D1	-10.00	-161.00
BME_01D0	32.37	-64.65	POCS05_01D1	-5.00	-159.00
BMW_01D0	32.27	-64.88	POC000_01D1	0.00	-155.00
BNE030_01P2	40.80	-97.18	POCN05_01D1	5.00	-151.00
BNE050_01P2	40.80	-97.18	POCN10_01D1	10.00	-149.00
BNE070_01P2	40.80	-97.18	POCN15_01D1	15.00	-145.00
BRW_01D0	71.32	-156.61	POCN20_01D1	20.00	-141.00
BRW_01C0	71.32	-156.61	POCN25_01D1	25.00	-139.00
CAR030_01P2	40.37	-104.30	POCN30_01D1	30.00	-135.00
CAR040_01P2	40.37	-104.30	POCN35_01D1	35.00	-137.00
CAR050_01P2	40.37	-104.30	POCN40_01D1	40.00	-136.00
CAR060_01P2	40.37	-104.30	POCN45_01D1	45.00	-131.00
CAR070_01P2	40.37	-104.30	PSA_01D0	-64.92	-64.00
CAR080_01P2	40.37	-104.30	RPB_01D0	13.17	-59.43
CBA_01D0	55.21	-162.72	RTA005_01P2	-21.25	-159.83
CFA_02D0	-19.28	147.06	RTA015_01P2	-21.25	-159.83
CGO_01D0	-40.68	144.69	RTA025_01P2	-21.25	-159.83
CHR_01D0	1.70	-157.17	RTA035_01P2	-21.25	-159.83
CMA030_01P2	38.83	-74.32	RTA045_01P2	-21.25	-159.83

CMA050_01P2	38.83	-74.32	RYO_19C0	39.03	141.83
CMA070_01P2	38.83	-74.32	SCA030_01P2	32.77	-79.55
CMO_01D0	45.48	-123.97	SCA050_01P2	32.77	-79.55
COI_20C0	43.15	145.50	SCA070_01P2	32.77	-79.55
CPT_36C0	-34.35	18.49	SCSN03_01D1	3.00	105.00
CRZ_01D0	-46.45	51.85	SCSN06_01D1	6.00	107.00
CSJ_06D0	51.93	-131.02	SCSN09_01D1	9.00	109.00
CYA_02D0	-66.28	110.52	SCSN12_01D1	12.00	111.00
DND030_01P2	48.38	-99.00	SCSN15_01D1	15.00	113.00
DND050_01P2	48.38	-99.00	SCSN18_01D1	18.00	113.50
DND070_01P2	48.38	-99.00	SCSN21_01D1	21.00	114.00
EIC_01D0	-27.15	-109.45	SEY_01D0	-4.67	55.17
ESP_02D0	49.58	-126.37	SGP015_01P2	36.80	-97.50
ESP005_01P2	49.58	-126.37	SGP025_01P2	36.80	-97.50
ESP015_01P2	49.58	-126.37	SGP035_01P2	36.80	-97.50
ESP025_01P2	49.58	-126.37	SGP045_01P2	36.80	-97.50
ESP035_01P2	49.58	-126.37	SHM_01D0	52.72	174.10
ESP045_01P2	49.58	-126.37	SIS_02D0	60.17	-1.17
ESP055_01P2	49.58	-126.37	SMO_01D0	-14.25	-170.56
ETL010_01P2	54.35	-104.98	SMO_01C0	-14.25	-170.56
ETL030_01P2	54.35	-104.98	SPLDTA_03C0	40.45	-106.73
ETL050_01P2	54.35	-104.98	SPO_01D0	-89.98	-24.80
ETL070_01P2	54.35	-104.98	SPO_01C0	-89.98	-24.80
GMI_01D0	13.43	144.78	STM_01D0	66.00	2.00
GOZ_01D0	36.05	14.18	STMEBC_01D0	66.00	2.00
GSN_24D0	33.28	126.15	STP_12D0	50.00	-145.00
HAA005_01P2	21.23	-158.95	SUM_01D0	72.58	-38.48
HAA015_01P2	21.23	-158.95	SYO_01D0	-69.00	39.58
HAA025_01P2	21.23	-158.95	TAP_01D0	36.73	126.13
HAA035_01P2	21.23	-158.95	TDF_01D0	-54.87	-68.48
HAA045_01P2	21.23	-158.95	TGC005_01P2	27.73	-96.86
HAA055_01P2	21.23	-158.95	TGC015_01P2	27.73	-96.86
HAA065_01P2	21.23	-158.95	TGC025_01P2	27.73	-96.86
HAA075_01P2	21.23	-158.95	TGC035_01P2	27.73	-96.86
HAT_20C0	24.05	123.80	TGC045_01P2	27.73	-96.86
HBA_01D0	-75.58	-26.50	TGC055_01P2	27.73	-96.86
HDPDTA_03C0	40.56	-111.65	TGC065_01P2	27.73	-96.86
HFM015_01P2	42.54	-72.17	TGC075_01P2	27.73	-96.86
HFM025_01P2	42.54	-72.17	THD005_01P2	41.05	-124.15
HFM035_01P2	42.54	-72.17	THD015_01P2	41.05	-124.15
HFM045_01P2	42.54	-72.17	THD025_01P2	41.05	-124.15
HFM055_01P2	42.54	-72.17	THD035_01P2	41.05	-124.15
HFM065_01P2	42.54	-72.17	THD045_01P2	41.05	-124.15
HFM075_01P2	42.54	-72.17	THD055_01P2	41.05	-124.15
HIL030_01P2	40.07	-87.91	THD065_01P2	41.05	-124.15
HIL050_01P2	40.07	-87.91	THD075_01P2	41.05	-124.15
HIL070_01P2	40.07	-87.91	TRM_11D0	-15.88	54.52
ICE_01D0	63.40	-20.29	UTA_01D0	39.90	-113.72

IZO_01D0	28.31	-16.50	UUM_01D0	44.45	111.10
IZO_27C0	28.31	-16.50	WBI030_01P2	41.72	-91.35
JBN_29C0	-62.23	-58.82	WBI050_01P2	41.72	-91.35
KEY_01D0	25.67	-80.16	WBI070_01P2	41.72	-91.35
KUM_01D0	19.52	-154.82	WKT030_01C3	31.31	-97.33
KZM_01D0	43.25	77.88	WKT122_01C3	31.31	-97.33
LEF010_01P2	45.95	-90.27	WKT457_01C3	31.31	-97.33
LEF020_01P2	45.95	-90.27	WPON30_20D2	30.00	146.00
LEF030_01P2	45.95	-90.27	WPON25_20D2	25.00	146.00
LEF040_01P2	45.95	-90.27	WPON20_20D2	20.00	146.00
LMP_01D0	35.52	12.62	WPON15_20D2	15.00	146.00
MAA_02D0	-67.62	62.87	WPON10_20D2	10.00	146.00
MBC_01D0	76.25	-119.35	WPON05_20D2	5.00	146.00
MHD_01D0	53.33	-9.90	WPO000_20D2	0.00	146.00
MHDCBC_11C0	53.33	-9.90	WPOS05_20D2	-5.00	146.00
MHDRBC_11C0	53.33	-9.90	WPOS10_20D2	-10.00	146.00
MID_01D0	28.21	-177.38	WPOS15_20D2	-15.00	146.00
MKN_01D0	-0.05	37.30	WPOS20_20D2	-20.00	146.00
MLO_01D0	19.54	-155.58	WPOS25_20D2	-25.00	146.00
MLO_01C0	19.54	-155.58	YON_19C0	24.47	123.02
MNM_19C0	24.30	153.97	ZEP_01D0	78.90	11.88
MQA_02D0	-54.48	158.97	ZOT015_45D2	60.75	89.38
NHA015_01P2	42.95	-70.63	ZOT025_45D2	60.75	89.38
NHA025_01P2	42.95	-70.63	ZOT035_45D2	60.75	89.38

Table 2.2. Monthly breakdown of the number of 5°×5° monthly-mean GOSAT X_{CO2} retrievals used in the one-year flux estimation.

Year/Month	GOSAT 5°×5°	GOSAT 5°×5° land	GOSAT 5°×5° ocean	Latitude of northern-most GOSAT data	Latitude of southern-most GOSAT data
0906	471	e	e	72.5	-47.5
0907	447	306	141	72.5	-47.5
0908	460	329	131	72.5	-47.5
0909	499	353	146	67.5	-47.5
0910	491	302	189	57.5	-47.5
0911	474	263	211	47.5	-42.5
0912	411	208	203	42.5	-47.5
1001	413	199	214	47.5	-42.5
1002	347	190	157	47.5	-47.5
e	390	227	163	52.5	-52.5
1004	390	241	149	62.5	-52.5
1005	425	278	147	67.5	-42.5
1006	420	318	102	82.5	-42.5
e	487	340	147	77.5	-42.5
Average	438	277	160		
Total	6125	3883	2242		

Figures

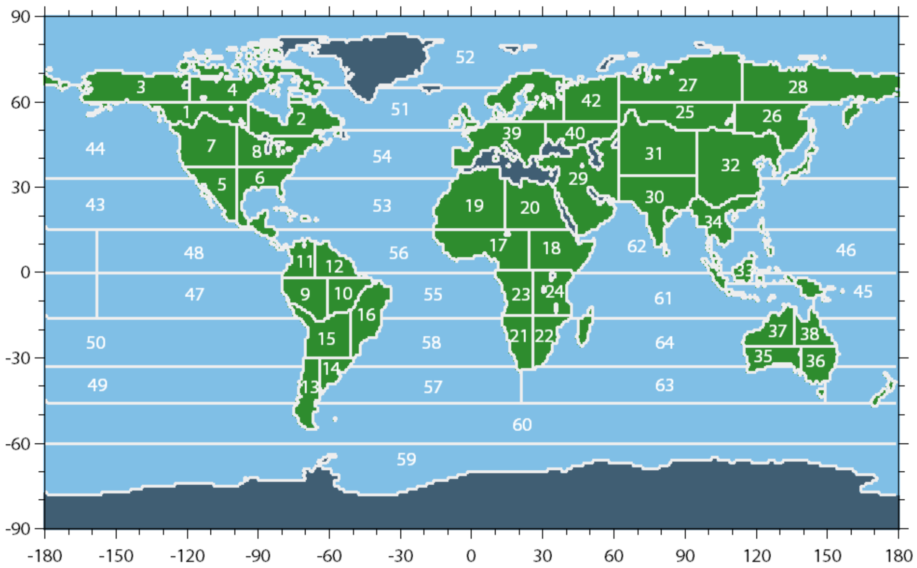

Figure 2.1. Boundaries of the 64 regions adopted in this study. The numbers on the figure are the region IDs. Regions shaded with dark blue are not considered in the flux estimation.

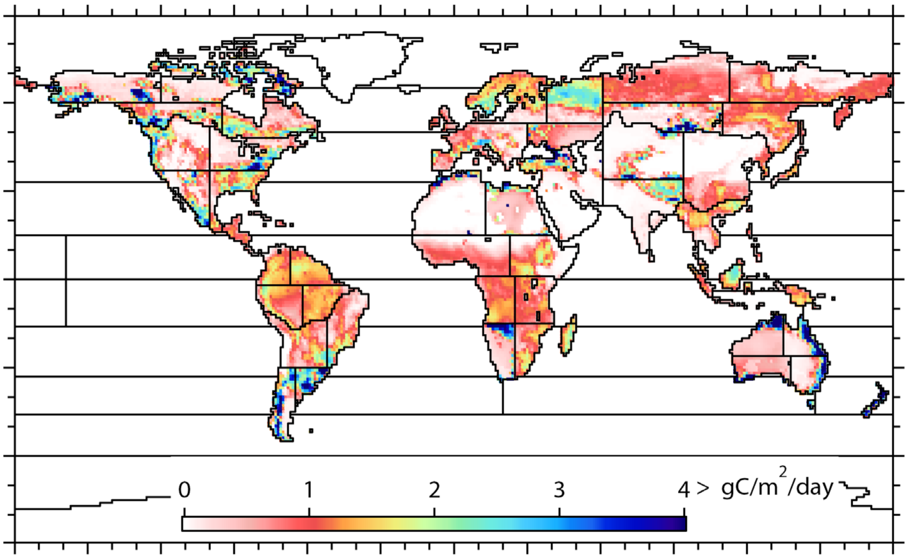

Figure 2.2. 1 Gt yr^{-1} $region^{-1}$ unit emission patterns for the 42 terrestrial regions. These spatial patterns were defined as that of 31-yr-mean net primary productivity estimated by VISIT (1980-2010).

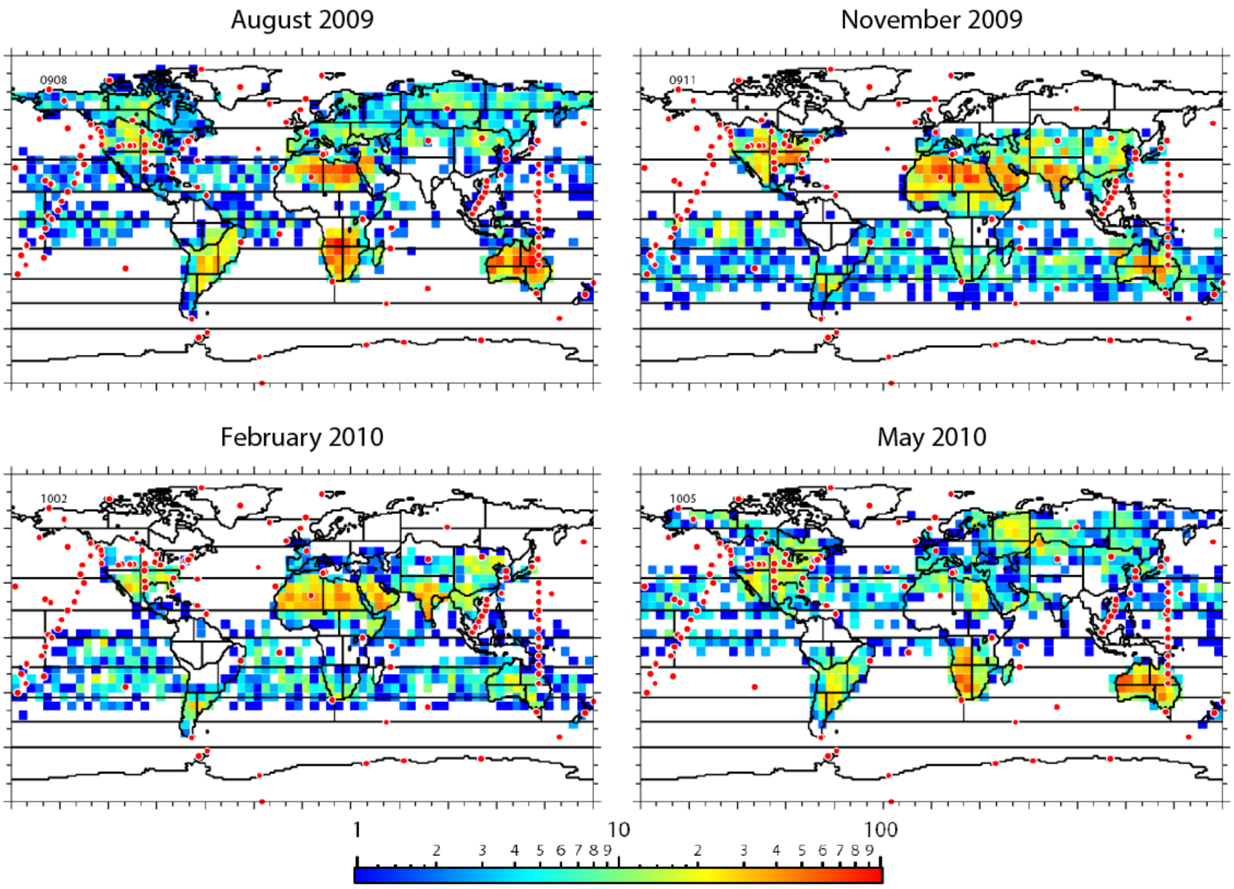

Figure 2.3. The number of GOSAT Level 2 X_{CO2} data records per each of 5°×5° grid cells during the months of August 2009, November 2009, February 2010, and May 2010. Red circles indicate the locations of the GV measurement sites chosen for this study.

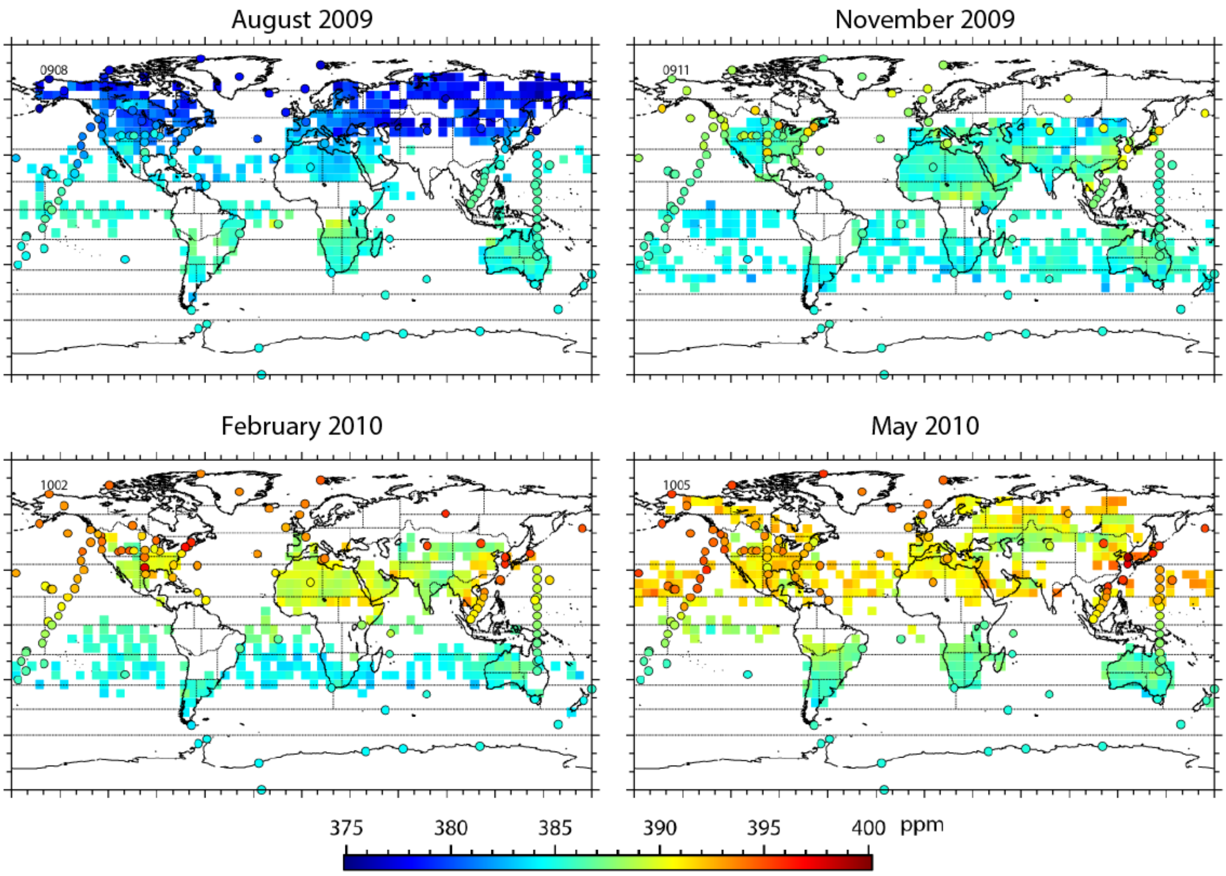

Figure 2.4. GOSAT X_{CO2} retrievals in the form of input to the inverse modeling scheme (gridded to 5°×5° cells and averaged on a monthly time scale). Cells with three or more X_{CO2} retrievals per month are shown. The bias was corrected by raising each X_{CO2} retrieval by 1.20 ppm. Overlaid are GLOBALVIEW values (in circles; monthly means). Values for the months of August 2009 (summer in the Northern Hemisphere), November 2009 (fall), February 2010 (winter), and May 2010 (spring) are shown.

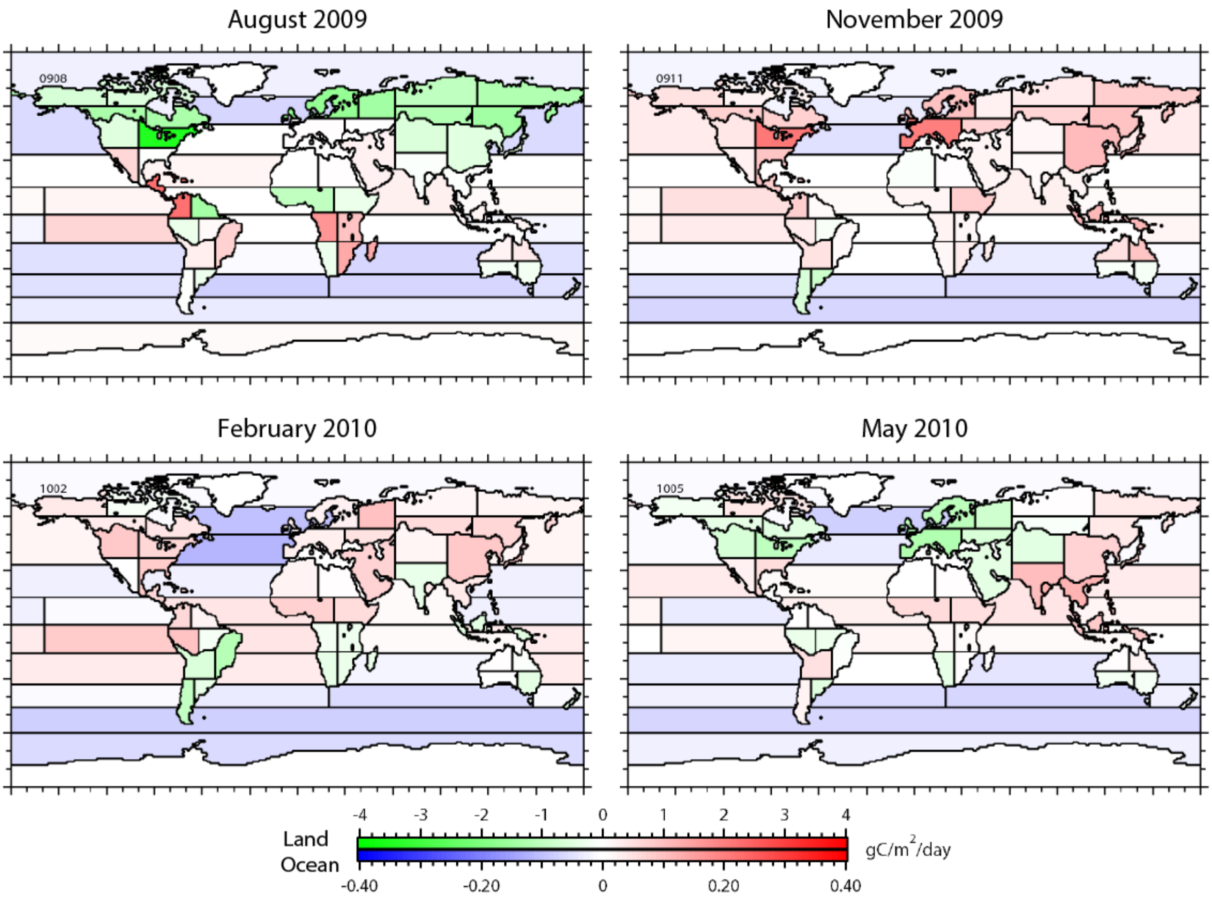

Figure 2.5. Monthly fluxes (g C m^{-2} day^{-1}) estimated for the 64 subcontinental regions using GV data and GOSAT X_{CO2} retrievals. Results for the months of August 2009 (summer in the Northern Hemisphere), November 2009 (fall), February 2010 (winter), and May 2010 (spring) are shown.

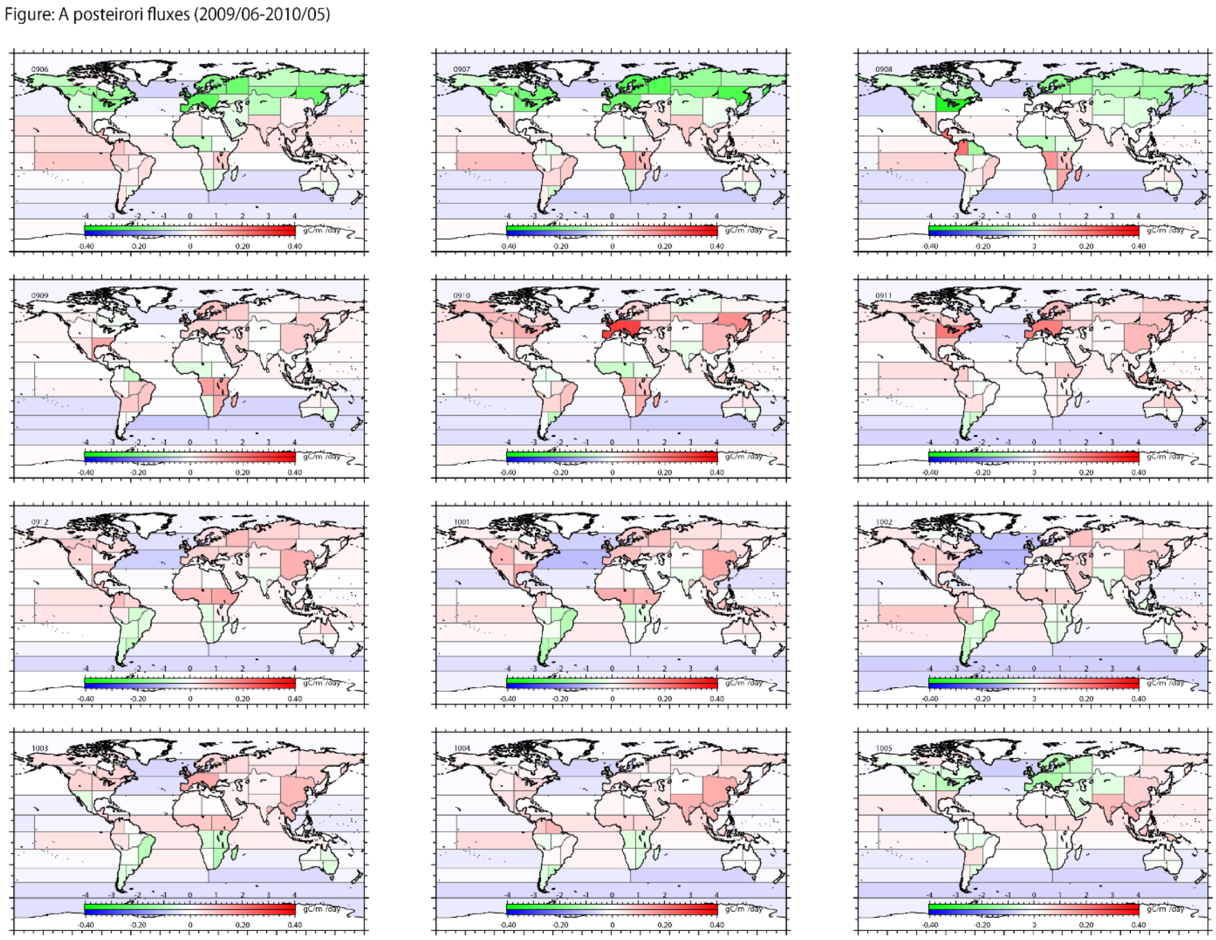

Figure 2.6. Monthly fluxes (g C m^{-2} day^{-1}) estimated for the 64 subcontinental regions using GV data and GOSAT X_{CO2} retrievals. Results for the analyzed one year are shown.

Figure: Concentration data used for inversion (2009/06-2010/05)

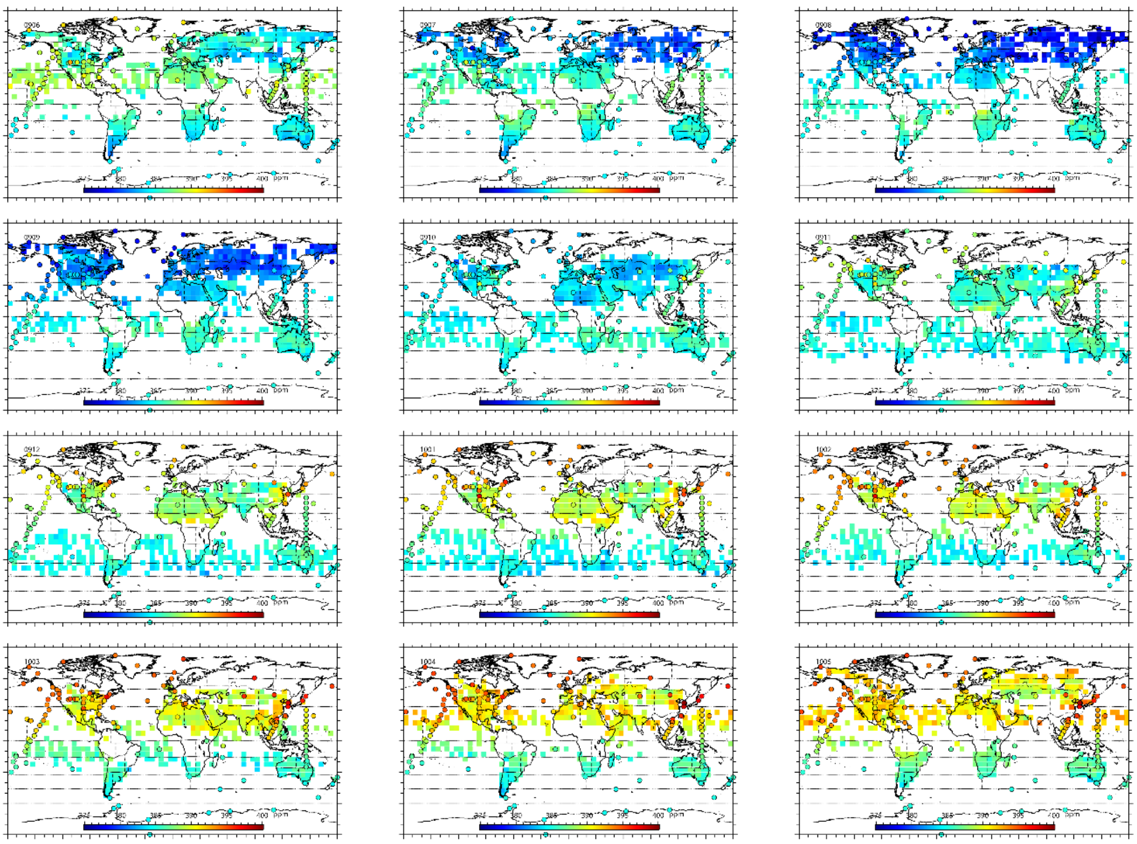

Figure 2.7. Monthly-mean concentration data used for the estimation of monthly fluxes presented in Figure 2.6 (12-months period).

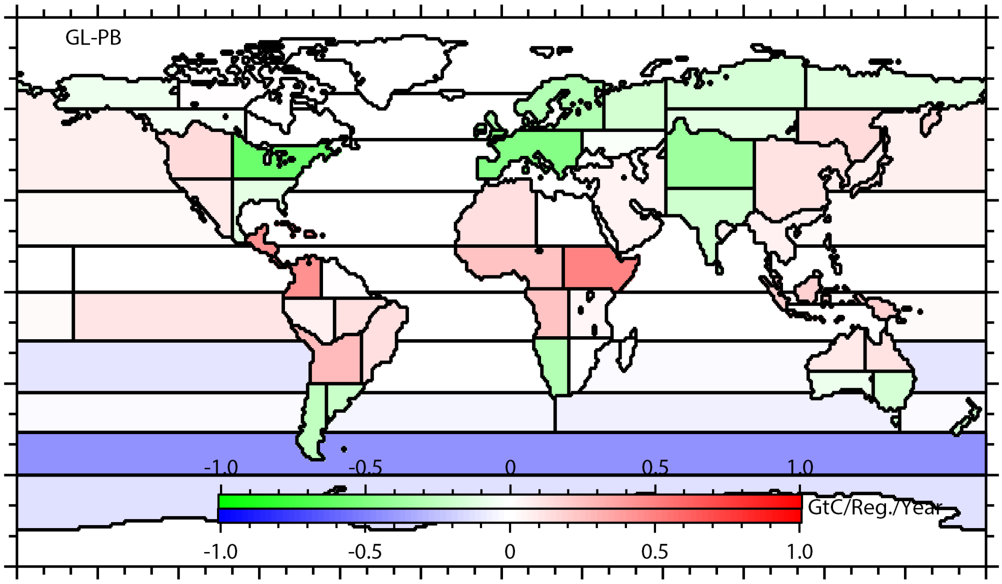

Figure 2.8. Annual-mean of the difference between the a posteriori and a priori fluxes (net). Values are shown as the a posteriori minus the a priori values (GtC region^{-1} year^{-1}).

CHAPTER 3

Utility of GOSAT data in regional monthly CO_2 flux estimation

This study was made possible through collaborating with the following researchers:

Robert J. Andres[1], Dmitry Belikov[2,3,4], Isamu Morino[2], Tomohiro Oda[5,6], Makoto Saito[2], Ryu Saito[2,*], Osamu Uchino[2], Vinu Valsala[7], Tatsuya Yokota[2], Yukio Yoshida[2], and Shamil Maksyutov[2]

1 Oak Ridge National Laboratory, TN, USA

2 National Institute for Environmental Studies, Tsukuba, Japan

3 National Institute of Polar Research, Tokyo, Japan

4 Tomsk State University, Tomsk, Russian Federation

5 Colorado State University, CO, USA

6 Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, CO, USA

7 Indian Institute for Tropical Meteorology, Pune, India

* Now at Kokusai Kogyo Co. Ltd., Tokyo, Japan

3.1. Introduction

Prior to the launch of GOSAT, Kadygrov et al. [2009], using an inversion system which was a predecessor to the one described in Chapter 2 and a set of pseudo GOSAT X_{CO2} retrievals, investigated the utility of GOSAT observations in the estimation of regional CO_2 fluxes. The dataset of the pseudo GOSAT X_{CO2} retrievals for 2005 was generated by running forward a set of climatological a priori fluxes using a version of NIES-TM [Maksyutov et al. 2008]. Clear-sky probabilities calculated from observational data collected in 2005 by the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) [Winker et al., 2006] were used in projecting the number distribution of successful X_{CO2} retrievals. The simulated X_{CO2} retrievals were then aggregated monthly to a 7.5°×7.5° grid. They concluded in this simulation analysis that the addition of the 7.5°×7.5° monthly GOSAT X_{CO2} retrievals with an assigned precision of 1.8 ppm to data from the existing surface monitoring sites (76 sites as used by Gurney et al. [2002]; see Figure 1.2 for the locations of the sites) can reduce the uncertainty of monthly regional surface fluxes as much as 50% (annual mean reduction).

Herein, I report the benefit of actual GOSAT observations to the estimation of CO_2 surface fluxes, using the established inversion system described in the previous chapter. For this, I estimated monthly regional fluxes and their uncertainty from 1) the 2011 issue of GV data [GLOBALVIEW-CO2 2011] and 2) both GV and version 02.00 of GOSAT Level 2 X_{CO2} retrievals, and compared these two sets of results in terms of flux uncertainty reduction as in the analysis by Kadygrov et al. [2009]. The rate of reduction in the flux uncertainty corresponds to the degree to which the GOSAT X_{CO2} retrievals contribute to constraining the surface fluxes. The analysis period is the one-year between June 2009 and May 2010, the first year of GOSAT sounding, as in Chapter 2.

3.2. Data and method

The GV data (220 sites) and 5°×5°-grid monthly mean GOSAT X_{CO2} retrievals, as well as the inverse modeling system, used for this analysis are the same as the ones described in Chapter 2. The results shown in here are thus based on the flux estimates obtained and described therein.

Shown in Figure 3.1 is the number of GOSAT X_{CO2} retrievals per each of 5°×5° grid cells over the one-year analyzed period. Overlaid onto this figure are the locations of the selected GV measurement sites whose data were used in this analysis (red circles; 220). Successful GOSAT X_{CO2} retrievals are particularly numerous over Africa, South America, and Australia, owing to the frequent occurrence of clear-sky days. For comparison, the number of pseudo GOSAT X_{CO2} retrievals in a 7.5°×7.5° grid for July 2005 as presented in the report by Kadygrov et al. [2009] is contrasted in Figure 3.2 with that of actual GOSAT retrievals obtained in July 2009 (on the same 7.5°×7.5° grid). The actual data number distribution shows high data volume over land in the Southern Hemisphere as the simulation indicates, but in the Northern Hemisphere, in particular the northwestern America and boreal Eurasia, it appears that the simulation may have over-predicted the successful retrievals. The difference can be also due to year-to-year variations in cloud cover distribtions.

3.3. Results

The reduction in the a priori flux uncertainty corresponds to the degree to which observations used in the flux inference contributed to determining, or "constraining", the surface fluxes. The reduction is often expressed by contrasting the diagonal parts of the a posteriori error covariance matrix, $\mathbf{C'_M}$, to that of the a priori one, $\mathbf{C_M}$. Here, it was rather chosen to consider the uncertainty reduction attained by the addition of the GOSAT X_{CO2} retrievals to the GV data. Following Rayner and O'Brian [2001], the uncertainty reduction (UR) in % is expressed as:

$$\mathrm{UR} = \left(1 - \frac{\sigma_{GV+GOSAT}}{\sigma_{GV}}\right) \times 100 \ ,$$

where σ_{GV} and $\sigma_{GV+GOSAT}$ denote the uncertainties in the monthly fluxes estimated from the GV data only and those from both the GV data and the GOSAT retrievals, respectively. For this evaluation, I implemented the inversion scheme using only the GV data to obtain flux estimates and the value of σ_{GV}. Figure 3.3 presents the UR values for August 2009, November 2009, February 2010, and May 2010. As indicated in Equation 2-12, the value of UR is affected by three factors: (1) the uncertainty in the observations and a priori fluxes, given by $\mathbf{C_D}$ and $\mathbf{C_M}$, respectively; (2) the sensitivity of observations to surface fluxes (determined by atmospheric transport and stored in **G**); and (3) the size of $\mathbf{C_D}$, which reflects the number of observations available for constraining the fluxes. Note that in the current inversion setup the uncertainties specified for GV data and that for GOSAT retrievals can differ by as much as one order of magnitude (e.g. the minimum uncertainty set for GV data and GOSAT retrievals is 0.3 and 3.0 ppm, respectively). This implies that the GV data have much greater weight in constraining regional fluxes. Also, there is approximately a one-order-of-magnitude difference between the uncertainties prescribed to land and ocean fluxes. These differences contribute to creating strong region-to-region or land-to-ocean contrasts in UR values, as seen in Figure 3.3.

Regions that are far from ground-based observation networks but are covered by GOSAT retrievals (e.g. Regions 29 (Temperate Asia SW) and 17 (Tropical Africa SW); see Figure 2.1 for identifying the regions) show higher UR values, with a maximum UR of 61% for region 29 in October 2009 (shown in Figure 3.4). However, the UR values for the North American and Australian regions (Regions 5–8 and 35–38) barely exceed ~15 %, despite the fact that GOSAT retrievals were constantly available within and around these regions throughout the 1-year analysis period (see Figure 2.3). This represents a case in which the constraint provided by the GV data prevails over that provided by the GOSAT X_{CO2} retrievals. Thus, higher URs in the figure highlight regions whose a posteriori fluxes were constrained by the GOSAT retrievals more strictly than those in other regions (Middle East, Asia, Africa, and South America). In light of the GOSAT mission objectives, Figure 3.3 indicates what the satellite was designed to perform in complementing the ground-based observations. However, care must be taken in evaluating the flux values, as these remote regions coincide with locations where the validation of GOSAT retrievals is not currently possible and the retrieval of X_{CO2} values itself is challenged by higher local surface albedo and/or contamination by clouds and aerosols.

Shown in Figure 3.5 are the annual means of monthly UR values over the June 2009-May 2010 analysis period. The uncertainty reductions attained over land on an annual basis ranged from 2 to 44 %; the mean UR over land was 10%. For comparison, the result of the annual uncertainty reduction analysis by Kadygrov et al. [2009] is shown in Figure 3.6. A dataset of pseudo GOSAT retrievals aggregated monthly on a 7.5°×7.5° grid and GV data from 76 sites, as opposed to 220 sites used in the present study, were used in their study. The commonalities found in these two annual estimates are that they both indicate low URs in temperate North America, Europe, and Australia, where a number of the GV stations exist. Also, the oceanic URs in both cases are very low (<5%). URs for temperate Asia, Africa, and mid-latitude South America in both cases, where GV data are sparse, are higher than those for regions with more GV data.

As implied in the differences between the number distributions of the pseudo and actual X_{CO2} retrievals shown in Figure 3.2, their result suggested that URs of up to about 40% can be attainable in boreal America and Eurasia, whereas the actual result turned out that the boreal URs are much lower than the expected (< ~15%). This may be attributed to the fact that the high latitudinal bands of the Northern Hemisphere during the winter months see nearly no X_{CO2} retrievals (Figure 2.3); the simulation may have been overestimating the available X_{CO2} retrievals there. Another contrast is found in the tropical South America regions (Regions 9-12). The actual result show that these regions received very small number of X_{CO2} retrievals there (Figure 3.1) and the annual URs for the regions were in the 5-15% range, whereas the simulation predicted much greater URs of over 40%.

Figure 3.7 shows the monthly time series of a priori flux (green line), a posteriori flux estimated from GV (red line), a posteriori flux estimated from both GV and GOSAT X_{CO2} retrievals (blue line), and the uncertainty reduction rate (gray vertical bar) for north-western Temperate North America (Region 7; top) and south-western Tropical Africa (Region 17; middle). Time series for the other terrestrial regions are found in Figure 3.8. Note here that the uncertainty reduction rate is variable in a year since the number of GOSAT X_{CO2} retrievals, which is subject to the occurrence of clear-sky days and the local solar zenith angle that affects the X_{CO2} retrieval, changes with season. Both regions received GOSAT X_{CO2} retrievals over the one-year period (>30 retrievals per grid within those regions; Figure 3.1), but these two regions are quite contrasting in the density of GV stations therein and nearby. This is clearly reflected in the difference in the flux uncertainty reduction. The flux inferred for north-western Temperate North America finds much less uncertainty reduction by GOSAT X_{CO2} retrievals than that for south-western Tropical Africa does. The trends of a posteriori fluxes estimated from GV only and GV and GOSAT X_{CO2} retrievals are nearly identical over the analysis period. This is attributed to the fact that the observation errors prescribed to GV data are nearly one order of magnitude smaller than those of GOSAT X_{CO2} retrievals (Sections 2.6.1 and 2.7), allowing GV data to constrain the flux more strictly than the GOSAT X_{CO2} retrievals do. New information brought by GOSAT is therefore found in the Tropical Africa a posteriori flux estimated from both GV and GOSAT X_{CO2} retrievals. Eastern Pacific South (Region

47; Figure 3.6 bottom) is one of the oceanic basins that received larger numbers of GOSAT X_{CO2} retrievals. The uncertainty reduction on the order of a few percent indicates the challenging nature of estimating oceanic fluxes, which are approximately one order of magnitude smaller than the terrestrial counterparts (see the ordinate of Figures 3.7 bottom for the flux scale; see also Figure 3.8), via the "top-down" Bayesian surface CO_2 flux inference.

3.4. Concluding Remarks

Here in this Chapter, UR was used as a metric to evaluate the degree of benefit the satellite-based X_{CO2} retrievals bring to the regional flux estimation. The GOSAT X_{CO2} retrievals were found to benefit the undersampled regions, such as Africa and Asia, most, reducing the a posteriori uncertainties as much as ~60%.

The results presented above were obtained by using the monthly means of the GV data records and GOSAT X_{CO2} retrievals gridded to 5°×5° cells. One important aspect to note here is that the reduction of a posteriori flux uncertainty is dependent on the number of the observations used for constraining surface fluxes. The number of observations available for constraining surface fluxes is significantly reduced via averaging (e.g., a few tens of observations in a grid cell down to a single monthly mean). Thus, the result presented herein shows only a portion of the full benefit that GOSAT soundings can bring to the surface CO_2 flux estimation.

References

GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide. NOAA ESRL, Boulder, Colorado, 2011.

Gurney, K. R., et al. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, *Nature*, **415**, 626–630.

Kadygrov, N., S. Maksyutov, N. Eguchi, T. Aoki, T. Nakazawa, T. Yokota, and G. Inoue (2009), Role of simulated GOSAT total column CO_2 observations in surface CO_2 flux uncertainty reduction, *J. Geophys. Res.*, **114**, D21208, doi:10.1029/2008JD011597.

Maksyutov, S., P. K. Patra, R. Onishi, T. Saeki, and T. Nakazawa (2008), NIES/FRCGC global atmospheric tracer transport model: Description, validation, and surface sources and sinks inversion, *J. Earth Simulator*, **9**, 3 –18.

Morino, I., et al., (2011) Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, *Atmos. Meas. Tech.*, **4**, 1061-1076, doi:10.5194/amt-4-1061-2011.

Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO_2 concentration data in surface source inversions. *Geophys. Res. Lett.*, **28**, 175−178.

Winker, D. M., C. A. Hostetler, M. A. Vaughan, and A. H. Omar (2006), CALIOP Algorithm Theoretical Basis Document, Part 1: CALIOP Instrument, and Algorithms Overview, NASA Tech. Rep., PC-SCI-202, part 1.

Figures

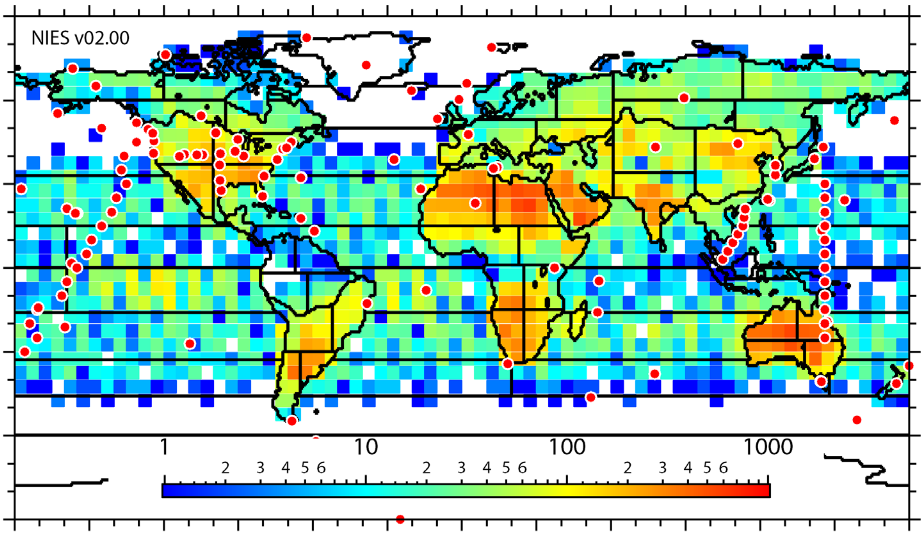

Figure 3.1. The number of GOSAT Level 2 X_{CO2} data per each of 5°×5° grid cells in a 12-months period between June 2009 and May 2010. The red circles indicate the locations of the GV measurement sites chosen for this study (220 sites).

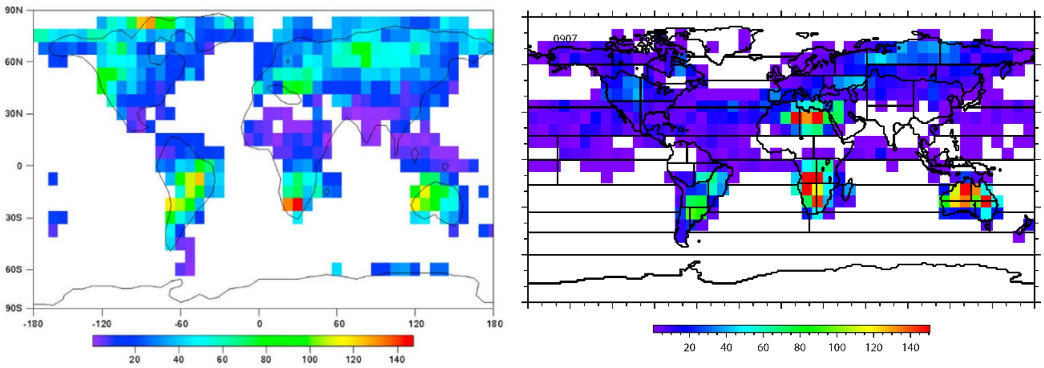

Figure 3.2. Left: the number of pseudo GOSAT X_{CO2} retrievals in a 7.5°×7.5° grid for July 2005 as presented in the report by Kadigrov et al. [2009]. Right: the number of successful GOSAT retrievals obtained in July 2009 (on the same 7.5°×7.5° grid).

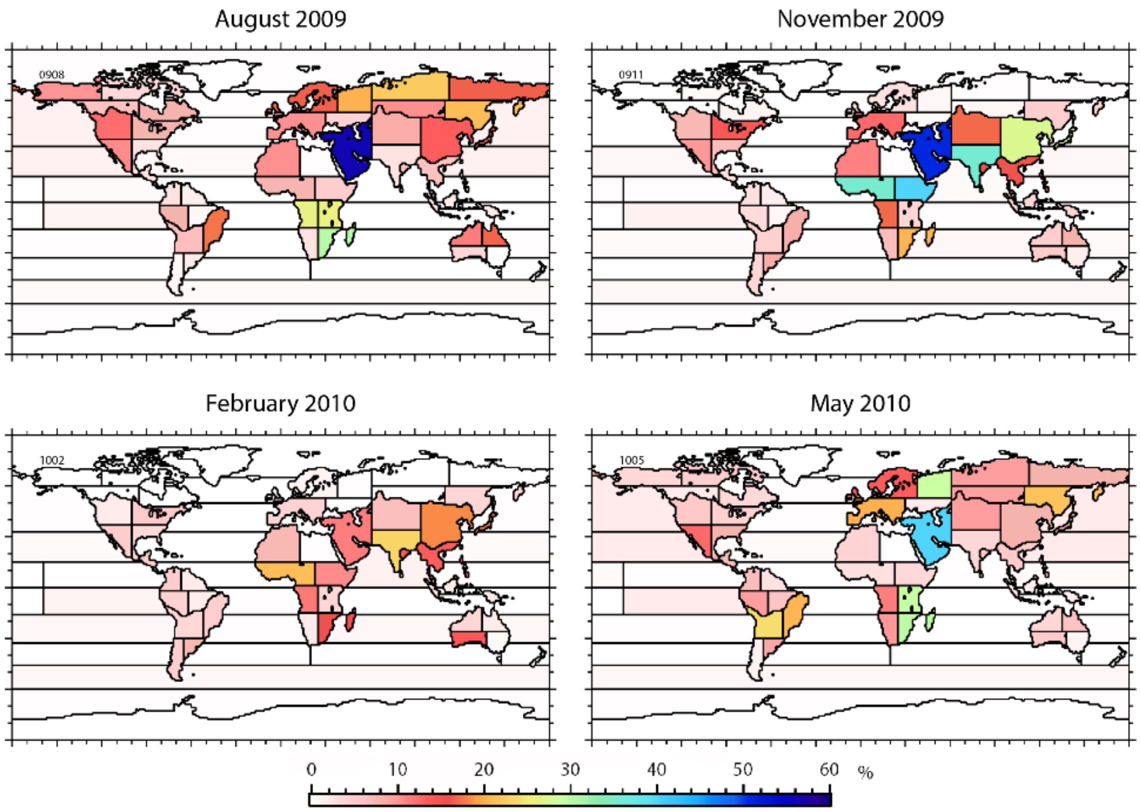

Figure 3.3. Percent reduction in the uncertainty of monthly surface flux estimates, attained by adding the GOSAT X_{CO2} retrievals to the GLOBALVIEW dataset.

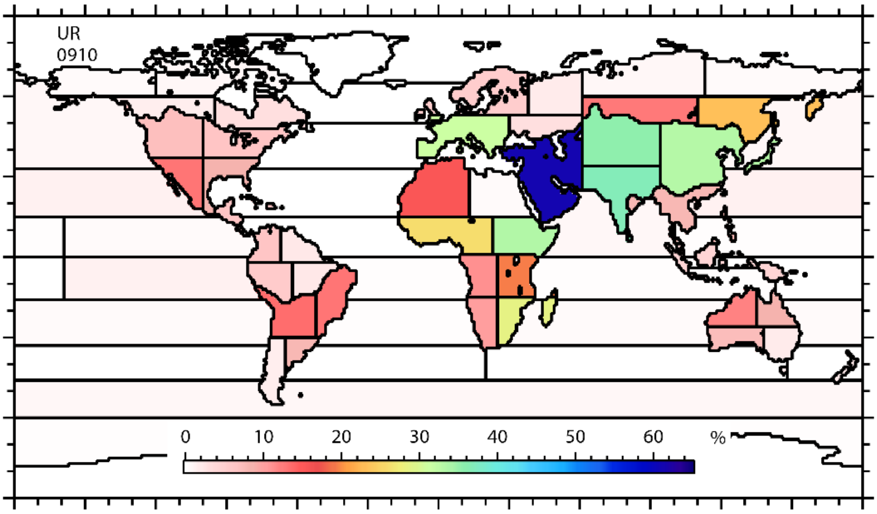

Figure 3.4. UR for October, 2009.

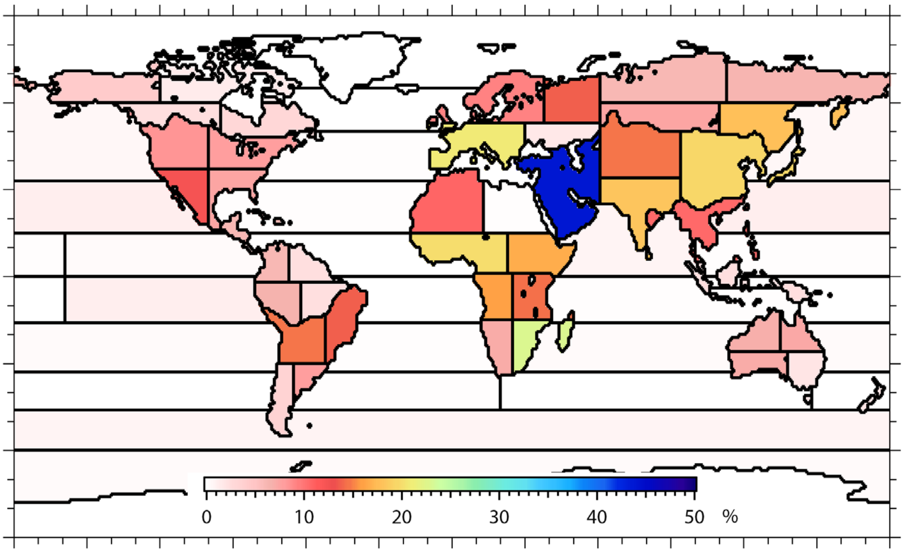

Figure 3.5. Annual mean URs over the June 2009-May 2010 analysis period.

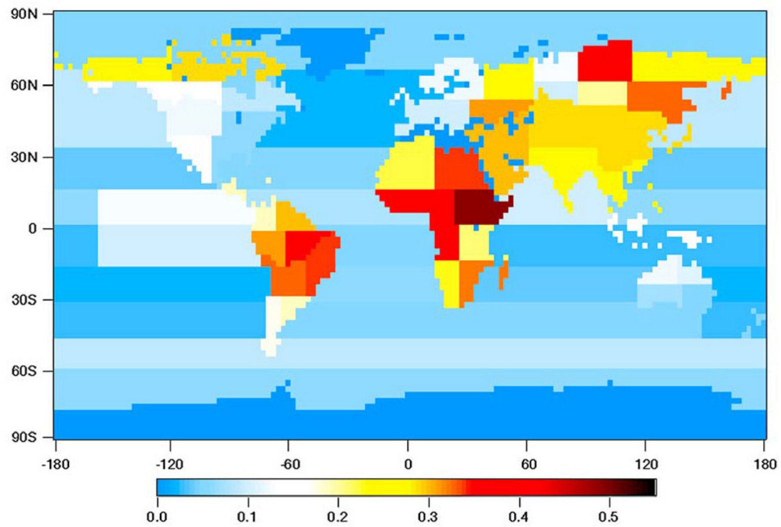

Figure 3.6. Annual uncertainty reduction (in fraction) predicted by Kadygrov et al. [2009] for the year 2005 (figure after Kadygrov et al., [2009]). A pseudo dataset of GOSAT X_{CO2} retrievals aggregated monthly to a 7.5° × 7.5° grid was used for the simulation. An uncertainty of 1.8 ppm was prescribed to each of the gridded X_{CO2} values used.

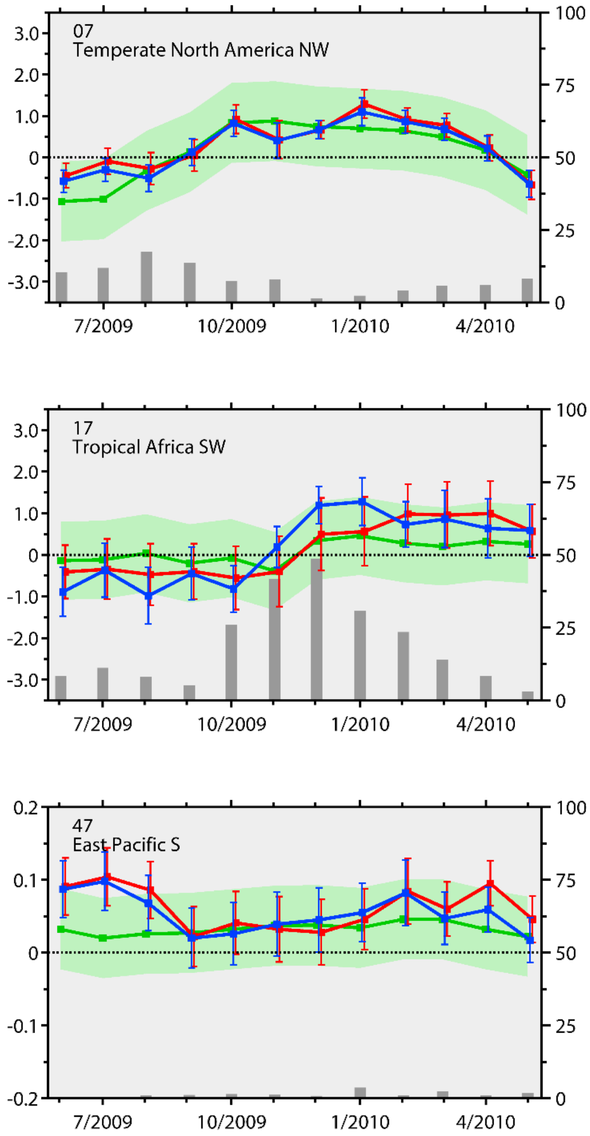

Figure 3.7. The time series of a priori flux (green), a posteriori fluxes estimated from GV (red), a posteriori flux estimated from both GV and GOSAT X_{CO2} retrievals (blue), and the uncertainty reduction rate (gray vertical bars). The blue shade indicates the a priori flux uncertainty. The error bar (red and blue) shows the a posteriori flux uncertainty. Results for north-western Temperate North America (Region 7; top panel), south-western Tropical Africa (Region 17; middle panel), and Eastern Pacific South (Region 47; bottom panel) are shown.

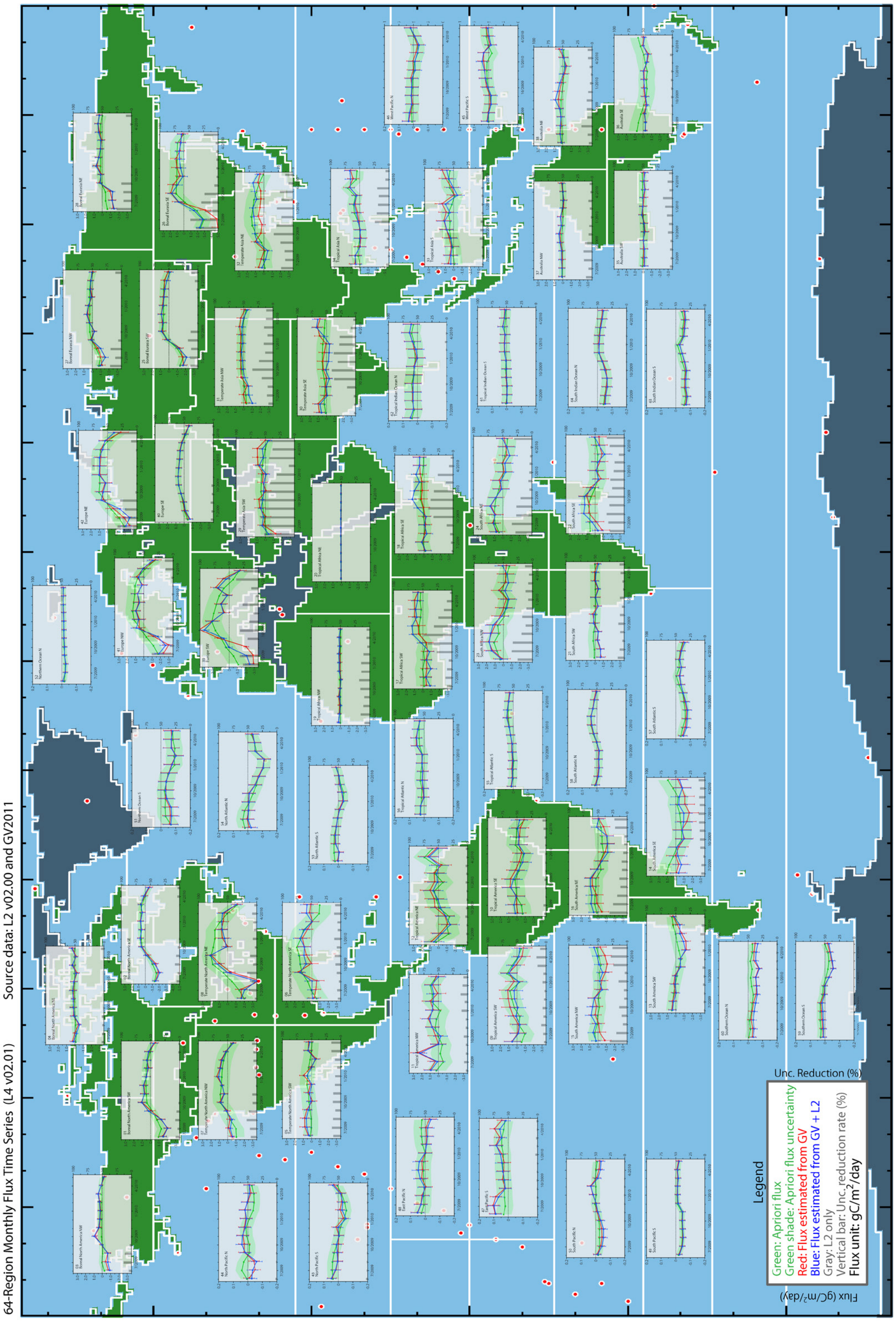

Figure 3.8. Time series of estimated fluxes, as shown in Figure 3.7, for the other remaining regions of the globe.

CHAPTER 4

Influence of differences in GOSAT X_{CO2} datasets on surface flux estimation

This study was made possible through collaborating with the following researchers:

Robert J. Andres[1], Dmitry Belikov[2,3,4], Andrey Bril[2], Hartmut Boesch[5], Andre Butz[6], Sandrine Guerlet[7,*] Otto Hasekamp[7], Sander Houweling[7,8], Isamu Morino[2], Tomohiro Oda[9,10], Christopher W. O'Dell[9], Sergey Oshchepkov[2], Robert Parker[5], Makoto Saito[2], Osamu Uchino[2], Vinu Valsala[11], Tatsuya Yokota[2], Yukio Yoshida[2], and Shamil Maksyutov[2]

1 Oak Ridge National Laboratory, TN, USA

2 National Institute for Environmental Studies, Tsukuba, Japan

3 National Institute of Polar Research, Tokyo, Japan

4 Tomsk State University, Tomsk, Russian Federation

5 EOS Group, Department of Physics and Astronomy, University of Leicester, Leicester, UK

6 Karlsruhe Institute of Technology, Leopoldshafen, Germany

7 Netherland Institute for Space Research, Utrecht, The Netherlands

8 Institute of Marine and Atmospheric Research Utrecht, Utrecht, The Netherlands

9 Colorado State University, CO, USA

10 Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, CO, USA

11 Indian Institute for Tropical Meteorology, Pune, India

* Now at Laboratoire de Météorologie Dynamique/IPSL, Paris, France

4.1. Introduction

The history of retrieving X_{CO2} from satellite-based SWIR spectral soundings traces back only to the period after the launch of SCIAMACHY (SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY) instrument aboard the European ENVISAT [Bovensmann et al., 1999] in 2002. Reports on initial SCIAMACHY-based X_{CO2} retrievals and algorithm development work were made by Buchwitz et al. [2005] and Barkley et al. [2006]. Later, these pioneering attempts were followed by the efforts of four independent groups that were involved in the research work of retrieving X_{CO2} from measurements by GOSAT, which was launched after SCIAMACHY/ENVISAT in 2009 to collect high-precision spectral soundings. In the GOSAT research community, there exist, as of 2013, five retrieval algorithms developed by the four groups: the National Institute for Environmental Studies (NIES), Japan (NIES v02 and PPDF-S) [Yoshida et al., 2013; Oshchepkov et al., 2013a], the NASA Atmospheric CO_2 Observations from Space (ACOS) team (ACOS B2.10) [O'Dell et al., 2012], the Netherlands Institute for Space Research / Karlsruhe Institute of Technology, Germany (RemoTeC v2.0) [Butz et al., 2011; Guerlet et al., 2013], and University of Leicester, UK (UoL-FP v3G) [Boesch et al., 2011; Cogan et al., 2012]. These algorithms have already gone through several updates since the launch of GOSAT. Although the algorithm improvement efforts continue, recent comparisons of the five X_{CO2} retrievals to the ground-based TCCON reference data showed that the mean and standard deviation of the GOSAT-TCCON differences are on the order of a few tenths of a percent [e.g.

Oshchepkov et al., 2013b]. With this progress, the first attempts at estimating CO_2 fluxes from the GOSAT-based X_{CO2} retrievals were made by multiple inverse modeling groups, and the results were cross-compared in the GOSAT CO_2 inversion inter-comparison campaign [Houwelling et al., in review]. The goal of the study is to assess the range of differences and the benefit of using GOSAT-based X_{CO2} retrievals in the flux estimation. In the initial stage of the campaign, each group used their choice of inverse modeling scheme and X_{CO2} retrieval dataset in obtaining their flux estimates. The result of the first assessment, focused on a one-year period from June 2009 to May 2010, are reported by Houwelling et al. [in review].

For evaluating and characterizing differences in flux estimates that are based on various modeling setups and concentration datasets, it is critical to know individual contributions from 1) the inverse modeling systems and 2) the X_{CO2} retrievals. I herein present the result of the latter assessment, which was obtained by estimating CO_2 fluxes from the five different X_{CO2} retrieval datasets using a single inverse modeling system, for the same one year between June 2009 and May 2010.

4.2. Data and method

4.2.1 Differences in X_{CO2} retrievals

The flow of data processing common to all of the five X_{CO2} retrieval algorithms is as follows: 1) pre-screening of GOSAT Level 1B SWIR spectral radiance data for perturbations by clouds and aerosols, 2) simulating the measured radiance spectra with a forward radiative transfer model, 3) retrieving X_{CO2} by optimizing the fit to the observed spectra, and 4) post-screening for low-quality X_{CO2} retrievals. The details of the implementation of these steps vary among the individual retrieval algorithms. Some key differences among the algorithms, as well as the number of successful land X_{CO2} retrievals yielded by each algorithm over the analyzed period, are shown in the upper part of Table 4.1.

The assessment of biases in the obtained X_{CO2} values, as discussed in Section 2.7.2, is an integral part of post-retrieval data validation. The lower part of Table 4.1 lists global-mean GOSAT-TCCON differences of the five retrieval datasets. Results based on both bias-corrected and uncorrected datasets (in parentheses) are shown. Biases in PPDF-S, ACOS, RemoTeC, and UoL-FP datasets were analyzed and corrected using multivariate linear regressions with which variabilities in X_{CO2} values were correlated with retrieval parameters such as surface albedo. The regression-based bias analysis for the NIES dataset (v02.00) was underway at the start of the GOSAT CO_2 inversion inter-comparison campaign, and for the current study the bias was corrected by raising each retrieved value by a global-mean GOSAT-TCCON difference (1.2 ppm). While debates on how to best analyze and correct biases outside the TCCON sites still continue, efforts are also devoted to investigating the causes of the biases. For more detailed descriptions on each of the five algorithms, including the bias correction approaches adopted, I refer the readers to a report on GOSAT retrieval algorithm inter-comparison by Oshchepkov et al. [2013b] and literature listed in Table 4.1.

Figure 4.1 shows the standard deviations (SD) of collocated X_{CO2} retrievals by the five algorithms for July 2009. The left panel shows SDs of coincident X_{CO2} retrievals to which bias corrections were applied, and the right panel presents those of uncorrected retrievals. Note that the geographical distribution of these coincidences does not represent that of any particular retrieval dataset (see Figure 4.2 for the distributions of five X_{CO2} datasets for July 2009). Only a fraction of five X_{CO2} datasets was found to coincide (see Figure 4.3 for coincidences in other months in the analyzed one year), thus values on these figures do not represent the spatial coverage of the individual datasets. Yet, Figure 4.3 indicates that the application of bias correction diminishes the spread among the five retrievals over the analyzed one year period. The global-mean SDs of the bias-corrected and uncorrected retrievals for July 2009 were 1.2 and 1.8 ppm, respectively. Over the whole analysis period, the global-mean SDs turned out to be 1.2 ppm (min.: 0.2; max.: 4.5) and 1.6 ppm (min.: 0.2; max.: 5.4), respectively (Table 4.2 A and B show monthly statistics). Despite that the bias correction reduced the global-mean biases to nearly zero (Table 4.1), SDs of GOSAT-TCCON differences, both before and after the application of bias correction, remain approximately 2 ppm. The GOSAT-TCCON difference SDs, shown in Figure 4.4, may suggest Gaussian distributions. This 2 ppm uncertainty was considered as a random error associated with the current versions of X_{CO2} retrieval datasets, and it was taken into account in the flux estimation as the GOSAT data uncertainty (described in the next section).

4.2.2 Experimental setup

The inversion system described in Chapter 2 was used in this experiment. The a priori flux data used here consist of ODIAC fossil fuel emissions (ver. 3), GFED biomass burning emissions (ver. 3.1), VISIT-simulated terrestrial biosphere NEE (ver. 3.0), and OTTM assimilated ocean-atmosphere exchange. Monthly regional fluxes and their uncertainties were estimated from each of the five X_{CO2} retrieval datasets that were combined with the 2011 issue of GV surface-based network data [GLOBALVIEW-CO2, 2011]. Data from 220 surface monitoring locations, including airborne sites, were used (see upper left panel of Figure 4.2 for locations). Following Law et al. [2003], the locations of all coastal sites used were shifted offshore in order to account for the selective measurements reflected in GV data. After performing the forward concentration simulation of each GV and X_{CO2} value, the GV values were monthly-averaged, and the X_{CO2} retrievals were gridded to 5°×5° cells and averaged on a monthly basis. The X_{CO2} retrievals were regularized this way to reduce the potential influence of differences in the number of X_{CO2} retrievals each algorithm yields (Table 4.1; the maximum difference is as large as ~40000 retrievals yr^{-1}) and in their horizontal coverage (Figure 4.2) on the flux estimation as much as possible. 5°×5° cells with less than three X_{CO2} retrievals per month were not considered here. The uncertainties for the GV values were taken from residual SDs about smooth curves that are stored in the GV 2011 dataset, and those for the X_{CO2} retrievals were determined as SDs of X_{CO2} retrievals found in each of 5°×5° grid cells in a month (all-data mean SD: 1.6 ppm; range: 0.02-7.8 ppm). Figures 4S.1-4S.5 show SD distributions for the five X_{CO2} datasets.

Following Law et al. [2003], I took account for errors associated with both the measurement and the forward concentration simulation by setting minimum uncertainties for the GV and X_{CO2} values at 0.3 and 3.0 ppm, respectively. The minimum uncertainty for X_{CO2} retrievals is based on the above-mentioned uncertainty associated with X_{CO2} retrieval (2.0 ppm) and error in the simulation of vertical column concentrations (~1.0 ppm) as reported by Belikov et al. [2013].

4.3. Results

4.3.1 Spread of five estimated fluxes due to differences in X_{CO2}

Presented in panels A and B of Figure 4.5 are the mean and SD of the five independent monthly fluxes for July 2009 estimated from the bias-corrected X_{CO2} retrievals. The fluxes shown include anthropogenic emissions. The influence of the X_{CO2} retrievals on these regional flux estimates is not uniform, but depends, among other factors, on the availability of both X_{CO2} retrievals and GV data within and around each region. To identify flux estimates on which X_{CO2} retrievals had large influence, I show in panel C the uncertainty reduction rate (UR) that represents the degree to which X_{CO2} retrievals contribute to constraining regional fluxes. As defined in Chapter 3, UR in percent is given as

$$\mathrm{UR}(\%) = \left(1 - \frac{\sigma_{GV+GOSAT}}{\sigma_{GV}}\right) \times 100,$$

where σ_{GV} and $\sigma_{GV+GOSAT}$ denote the uncertainties of fluxes estimated from the GV data alone and both the GV and X_{CO2} retrievals, respectively. Panel C shows the mean of five UR values. To distinguish cases with pronounced influence by GOSAT retrievals from those in ambiguity, I set a threshold of 10% UR, which comes from doubling the annual-mean URs of Amazonian regions (Regions 9 to 12) whose fluxes were constrained by data collected in distant regions since both GV data and X_{CO2} retrievals were nearly not present in these regions throughout the analyzed year. In panel B, terrestrial regions with URs greater than the threshold are indicated with asterisks. The statistical consistency of these above-UR-threshold GV+X_{CO2} fluxes with the corresponding GV-only values, which determines whether the GV-GOSAT joint estimation is a refinement of the GV-only case, is ensured by the fact that among the high-UR GV+X_{CO2} fluxes (total of 767 monthly estimates in the analyzed year; five flux datasets total), 93% of them were found within the uncertainty ranges (flux estimated ± a posteriori uncertainty) of the corresponding GV-only values, and in the remaining cases (7%), their uncertainty ranges overlapped those of the corresponding GV-only values.

Flux SDs for these high-UR regions ranged from 0.2 (Region 18) to 0.6 (Region 39) gC m^{-2} day^{-1}, and each of these SD values was found to be nearly equal or smaller than the mean of the corresponding a posteriori flux uncertainties (panel D). In the case of Region 39 (Europe SW; associated with the largest SD in the analyzed period), the spread between the largest and smallest flux estimates among the five results was 1.2 gC m^{-2} day^{-1}, which translated into a maximum SD of five a posteriori concentrations of 3.7 ppm (panel E; SD of monthly-mean concentrations simulated on a 2.5°×2.5° grid at 0.975 sigma level within Region 39).

(Figures for the other months of the analysis year are found in Figures 4S.5 –4S.15.)

4.3.2 Annual mean fluxes

To investigate the larger-scale influence of the differences in the five X_{CO2} retrievals on the flux estimation, I calculated annual global mean fluxes (net) and land/ocean partitions (without anthropogenic emissions) for each of the five inversion results. The values were obtained by aggregating the monthly regional fluxes, and are listed in Table 4.3 (unit: GtC yr^{-1}). The mean of the five annual global land uptakes was

1.7±0.3 GtC yr^{-1}. Relative to the GV-only result, all five results show reduction in global terrestrial biosphere uptake or enhancement in respiration.

To further explore this commonality, I show in Figure 4.6 annual regional fluxes estimated from GV data alone (panel A) and the mean of five GV+X_{CO2} annual regional fluxes (panel B). The anthropogenic and biomass burning emissions are not included here. Panel C shows the mean and SD of the departure of each of the annual mean GV+X_{CO2} estimates from the GV-only result. The values are shown as GV+X_{CO2} minus GV-only result. Similar to the approach presented in the previous section, annual regional flux estimates with pronounced influence of GOSAT retrievals were identified based on annual-mean UR values (mean URs ≥ 10%). Those are marked with asterisks in panel B and colored in panel C. URs of temperate North America (Regions 05-08) and Australia regions (Regions 35-38) were below the threshold because the fluxes were constrained more strongly by surface-based data because of their uncertainties that are smaller than those of X_{CO2} retrievals. URs of upper boreal regions (>~60°N) were low because GOSAT retrievals were only available during the local summer months. Oceanic URs were all below the threshold, and therefore only the terrestrial results are presented in panel C.

Integrated over the 11 continental-scale TransCom terrestrial regions (the names of the 11 regions are listed at the bottom of panel C), the GV-only annual estimates on panel A shows a pattern of tropical land regions (tropical America, tropical Africa, and tropical Asia) being CO_2 sources and Northern Hemisphere extra-tropics (temperate North America, Europe, and boreal Eurasia) being CO_2 sinks, which agrees with the results of surface-based, long-term inversion studies previously reported [Baker et al., 2006; Gurney et al., 2008; and Bruhwiler et al., 2011]. The GV+X_{CO2} result on panel B shows the same pattern, but in the finer 42 terrestrial-region sub-continental-scale framework (panel C), it indicates uptake reductions or respiration enhancements in northern parts of South America region (Regions 15 and 16), south eastern boreal Eurasia (Region 26), and north eastern temperate Asia (Region 32), which partly account for the changes of the global terrestrial uptake values from the GV-only result shown in Table 4.3. It also shows uptake enhancements or respiration reductions in northern parts of South Africa region (Regions 23 and 24), and south western temperate Asia (Region 30).

4.4. Discussion and concluding remarks

Among the departures of the high-UR GV+X_{CO2} flux estimates from the GV-only results presented in Figure 4.6 C (colored), values for Regions 16, 23, 24, and 26 are associated with small SDs (<0.1 GtC yr^{-1}), indicating that the flux estimates are less dependent on the choice of X_{CO2} dataset. The spatial coverage that each of the five 5°×5°-gridded X_{CO2} datasets shows over these regions was found to be similar to one another throughout the analyzed year (see Figures 4S.16 - 4S.26). The number of 5°×5°- gridded X_{CO2} data that cover Region 16 for July 2009, for instance, is nearly the same among the five datasets (8 to 9; see Figure 4.7 for the spatial coverage). On one hand, the departures for the remaining colored regions (15, 17, 18, 22, and 29 through 32) are variable with SDs greater than ~0.2 GtC yr^{-1}. The error bars of the values for Regions 18, 22, 29, and 31 cross the zero departure line in Figure 4.6 C, showing that the sign of the five departure values (enhancement or reduction) was not uniform in these cases. The larger SDs may be linked to the following: 1) the agreement among X_{CO2} retrievals within and around these regions, which did not appear on Figures 4.1 and 4.3, was difficult to reach, and/or 2) the horizontal distribution of the number of available X_{CO2} retrievals was quite different from dataset to dataset. While the former link remains to be unclear, the spatial coverage by each of the five 5°×5°-gridded X_{CO2} datasets was found to be different from one to another, particularly over the temperate Asia regions (see Figure 4.7). The number of 5°×5°-gridded data that cover Region 32 (temperate Asia NE) in July 2009, for instance, varied from 6 to 20, and that of individual X_{CO2} values (not averaged to monthly-gridded values) counted in the same region and month ranged from 57 to 161 (see Figure 4.7 for the distribution differences).

How strongly fluxes are constrained in the inversion (as reflected in UR values) depends on the number and geographical locations of the observations and the data uncertainty prescribed to them. The influence of differences in horizontal data coverage on a posteriori flux estimates has been addressed in previous surface-data-based inversion studies by Law et al [2003] and Bruhwiler et al. [2011]. The implication is that the impact of the differences in the number of X_{CO2} retrievals may be more pronounced if they were processed in the inversion without any application of data number regularization as in the present study. A check on the sensitivity of SDs of the departures (shown in Figure 4.6 C) to changes in the minimum uncertainty for the X_{CO2} retrievals reveals that with a reduction by 1 ppm (reduced from 3 to 2 ppm; meaning more constraint exerted by X_{CO2} retrievals), SDs of the temperate Asia departures increase by ~23%. Care should be taken in analyzing flux estimates of regions in which the number of X_{CO2} retrievals varies largely from dataset to dataset as in the case of Region 32.

References

Baker, D. F., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles,* **20**, GB1002, doi:10.1029/2004GB002439.

Belikov, D. A., et al. (2013), Simulations of column-averaged CO_2 and CH_4 using the NIES TM with a hybrid sigma-isentropic (σ-θ) vertical coordinate, *Atmos. Chem. Phys.*, **13**, 1713-1732, doi:10.5194/acp-13-1713-2013.

Boesch, H., D. Baker, B. Connor, D. Crisp and C. Miller (2011), Global characterization of CO_2 column retrievals from shortwave-infrared satellite observations of the Orbiting Carbon Observatory-2 Mission, *Remote Sens.*, **3**, 270-304; doi:10.3390/rs3020270.

Bruhwiler, L. M. P., A. M. Michalak, W. Peters, D. F. Baker, and P. Tans (2005), An improved Kalman Smoother for atmospheric inversions, *Atmos. Chem. Phys.*, **5**, 2691–2702, doi:10.5194/acp-5- 2691-2005.

Bruhwiler, L. M. P., Michalak, A. M., and Tans, P. P. (2011), Spatial and temporal resolution of carbon flux estimates for 1983–2002, *Biogeosciences*, **8**, 1309-1331, doi:10.5194/bg-8-1309-2011.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., N¨oel, S., Rozanov, V. V., Chance, K. V., and Goede, A. (1999), SCIAMACHY – mission objectives and measurement modes, *J. Atmos. Sci*., **56**, 127–150.

Buchwitz, M., de Beek, R., Noël, S., Burrows, J. P., Bovensmann, H., Bremer, H., Bergamaschi, P., Körner, S., and Heimann, M. (2005), Carbon monoxide, methane and carbon dioxide columns retrieved from SCIAMACHY by WFM-DOAS: year 2003 initial data set, *Atmos. Chem. Phys*., **5**, 3313–3329.

Butz, A., et al. (2011), Toward accurate CO_2 and CH_4 observations from GOSAT, *Geophys. Res. Lett.,* **38**, L14812, doi:10.1029/2011GL047888.

Cogan, A. J., et al. (2012), Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations, *J. Geophys. Res.*, **117**, D21301, doi:10.1029/2012JD018087.

GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide. NOAA ESRL, Boulder, Colorado, 2011.

Guerlet, S., et al. (2013), Impact of aerosol and thin cirrus on retrieving and validating X_{CO2} from GOSAT shortwave infrared measurements, *J. Geophys. Res. Atmos.*, **118**, 4887–4905, doi:10.1002/jgrd.50332.

Gurney, K. R., D. Baker, P. Rayner, and S. Denning (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.Houwelling, S., et al., An inter-comparison of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements, in review.

Law, R. M., Y.-H. Chen, K. R. Gurney and TransCom 3 Modelers (2003), TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, *Tellus* B, **55**, 580–595. doi: 10.1034/j.1600-0889.2003.00053.x.

O'Dell, C. W., et al. (2012), The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, **5**, 99-121, doi:10.5194/amt-5-99-2012.

Oshchepkov, S., et al. (2013a), Simultaneous retrieval of atmospheric CO_2 and light path modification from space-based spectroscopic observations of greenhouse gases: methodology and application to GOSAT measurements over TCCON sites, *Applied Optics*, **52**, 1339-1350.

Oshchepkov, S., et al. (2013b), Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space. Part 2: Algorithm intercomparison in the GOSAT data processing for CO_2 retrievals over TCCON sites, *J. Geophys. Res. Atmos.*, **118**, 1493–1512, doi:10.1002/jgrd.50146.Yoshida,Y., et al. (2013), Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data, *Atmos. Meas. Tech.*, **6**, 1533-1547, doi:10.5194/amt-6-1533-2013.

Tables

Table 4.1. Key differences in the five X_{CO2} retrieval algorithms

	NIES v02	**PPDF-S**	**ACOS B2.10[4]**	**RemoTeC v2.0**	**UoL-FP v3G[5]**
Number of vertical layers/levels	15 layers (fixed†)	22 levels (variable‡)	20 levels (fixed†)	12 layers (fixed†)	20 levels (variable‡)
Simultaneous retrieval of surface pressure	Yes	No (meteorological analysis data used)	Yes	No (meteorological analysis data used)	Yes
Cloud-contaminated data screening method (pre-screening)	CAI[1] Image data + 2μm-band radiance	CAI[1] Image data	Difference between retrieved surface pressure[2] and prior value	CAI[1] Image data + 2μm-band radiance	Difference between retrieved surface pressure[2] and prior value
Aerosol vertical distribution	6 layers	3 layers	20 layers	Normal distribution	Normal distribution
Types of aerosols/clouds modeled	4	-	4	1	3
Num. of land retrievals (1 yr: Jun. 2009- May 2010)	58933	65038	78529	39956	62067
Bias correction	Global uniform correction	Multivariate linear regression	Multivariate linear regression	Multivariate linear regression	Multivariate linear regression
Global mean and SD of GOSAT-TCCON difference[3] (ppm) (*before bias correction)	0.0±2.0 (-1.2±2.0)*	0.0±1.6 (0.1±1.8)*	-0.1±1.8 (-1.0±2.0)*	-0.1±2.1 (-2.3±2.2)*	-0.1±2.4 (0.2±2.6)*
Reference	Yoshida et al. 2013	Oshchepkov et al. 2013b	O'Dell et al. 2012 Wunch et al. 2011b	Butz et al. 2011 Guerlet et al. 2013	Boesch et al. 2011 Cogan et al. 2012

[1] CAI: Cloud and Aerosol Imager onboard GOSAT.

[2] Retrieved with an O_2 A-band-only algorithm based on an assumption of no clouds and aerosols present.

[3] Each X_{CO2} retrieval found within a ±2° grid box centered at each of 11 TCCON sites was compared with TCCON data that were averaged over ±30 min. of GOSAT overpass time. The 11 TCCON sites are Sodankyla (67.368°N, 26.663°E), Bialystok (53.230°N, 23.025°E), Bremen (53.104°N, 8.845°E), Orleans (47.970°N, 2.113°E), Garmisch (47.476°N,11.063°E), Park Falls (45.945°N, 90.273°W), Lamont (36.604°N, 97.486°W), Tsukuba (36.051°N, 140.122°E), Darwin (12.424°S,130.829°E), Wollongong (34.406°S, 150.879°E), and Lauder (45.038°S, 169.684°E).

[4] Only the retrievals based on GOSAT Level 1B spectral radiance data collected in high-gain mode (including oceanic retrievals) were used in this study.

[5] Only the terrestrial retrievals are available.

† Number of retrieval layers/levels are fixed (layer thickness or level varies with surface pressure).

‡ Number of retrieval levels varies with local surface pressure (only the number of the lowest few levels changes).

Table 4.2 A. The global-mean SDs of collocated X_{CO2} retrievals that were bias-corrected. Unit: ppm.

YYMM	Global mean SD	Minimum SD	Maximum SD	RANGE
0906	1.2	0.3	3.1	2.8
0907	1.2	0.3	3.2	3.0
0908	1.2	0.2	2.6	2.4
0909	1.2	0.2	2.8	2.7
0910	1.2	0.2	4.5	4.3
0911	1.1	0.2	2.7	2.4
0912	1.3	0.3	2.6	2.3
1001	1.3	0.3	3.1	2.8
1002	1.2	0.3	2.5	2.2
1003	1.0	0.2	2.8	2.5
1004	1.0	0.3	2.3	2.1
1005	1.1	0.3	2.6	2.4
Average	**1.2**			

Table 4.2 B. The global-mean SDs of collocated X_{CO2} retrievals whose biases were not corrected. Unit: ppm.

YYMM	Global mean SD	Minimum SD	Maximum SD	RANGE
0906	1.8	0.5	3.3	2.7
0907	1.8	0.6	4.0	3.4
0908	1.6	0.4	3.6	3.2
0909	1.8	0.4	3.6	3.2
0910	1.6	0.3	5.4	5.1
0911	1.5	0.4	3.1	2.6
0912	1.4	0.4	2.7	2.3
1001	1.5	0.3	3.7	3.4
1002	1.4	0.4	2.7	2.3
1003	1.5	0.4	3.6	3.2
1004	1.3	0.4	3.5	3.1
1005	1.4	0.2	3.1	3.0
Average	**1.6**			

Table 4.3. Annual mean fluxes in GtC yr^{-1} over the one-year analyzed period (Jun. 2009 – May 2010).

	GV only	NIES v02	PPDF-S	ACOS B2.10	RemoTeC v2.0	UoL-FP v3G	Mean and SD of five results
Global (net)	4.7	5.1	4.7	4.8	5.1	4.8	4.9±0.2
Land uptake*	2.2	1.3	2.1	1.8	1.4	1.8	1.7±0.3
Ocean uptake*	2.0	2.4	2.1	2.2	2.3	2.2	2.3±0.1

* (Uptake: absorption.) Land and ocean uptakes do not include anthropogenic emissions. Land uptakes include biomass burning emissions.

Figures

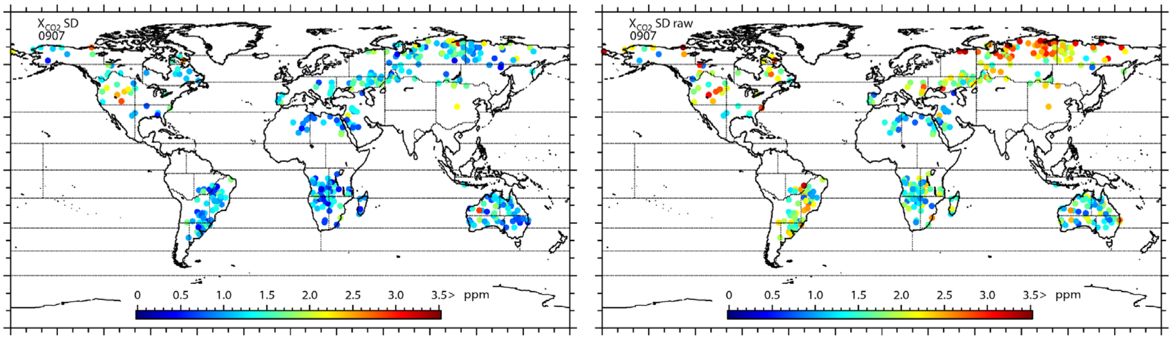

Figure 4.1. Standard deviation of five collocated X_{CO2} retrievals found in July 2009. Left: bias correction applied. Right: bias correction not applied.

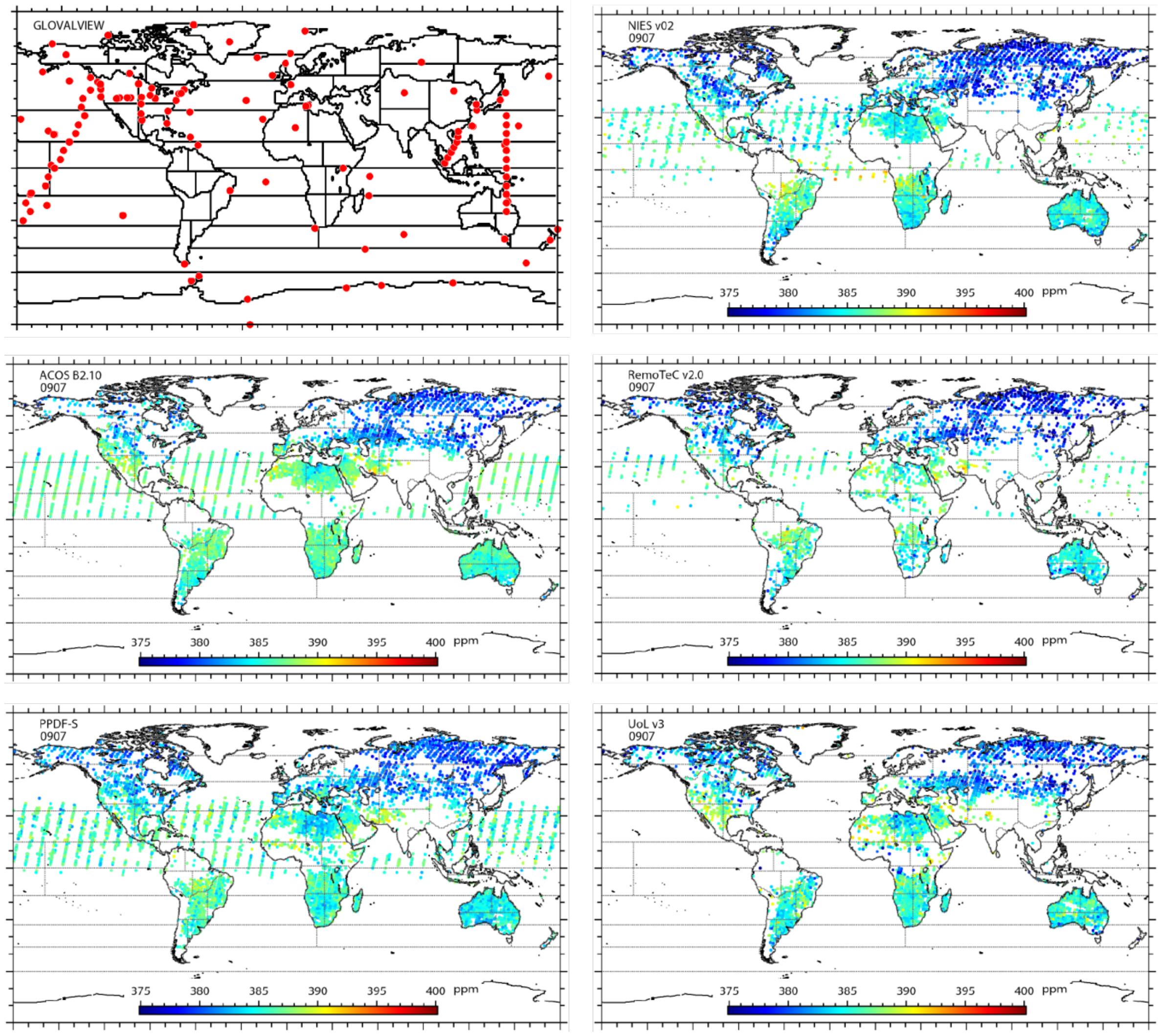

Figure 4.2. Distributions of five X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms. Bias corrections are applied. Values shown: July 2009. The upper left panel shows the locations of GLOVALVIEW-CO2 data sites selected for this analysis (220 stations including airborne sites).

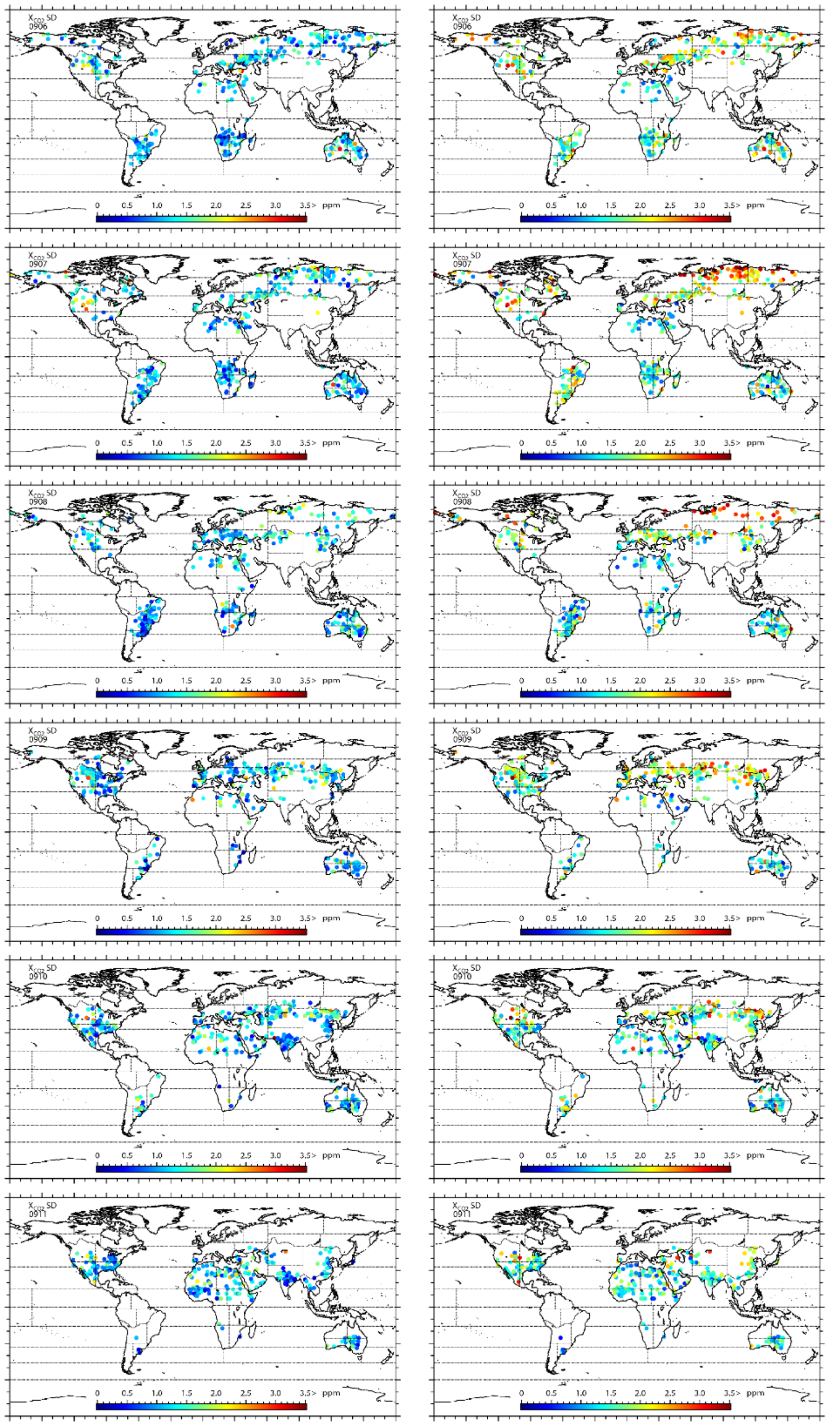

Figure 4.3. (1st half) SDs of five collocated X_{CO2} retrievals found in even-numbered months in the one-year analysis period. Top row: June 2009. Bottom row: November 2010. Left column: bias correction applied. Right column: no bias correction applied.

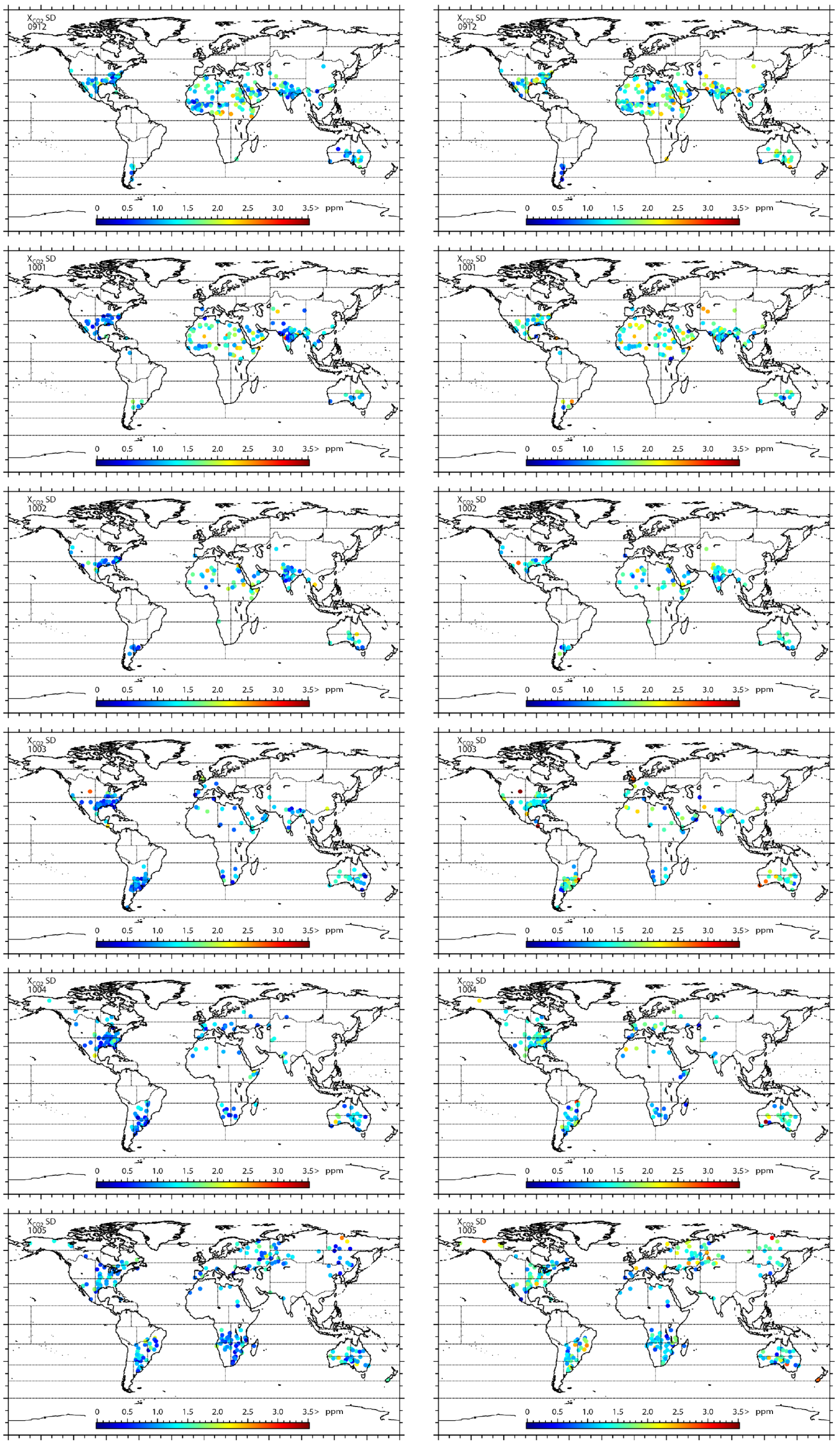

Figure 4.3. (2nd half) SDs of five coincident X_{CO2} retrievals found in even-numbered months in the one-year analysis period. Top row: December 2009. Bottom row: May 2010. Left column: bias correction applied. Right column: no bias correction applied.

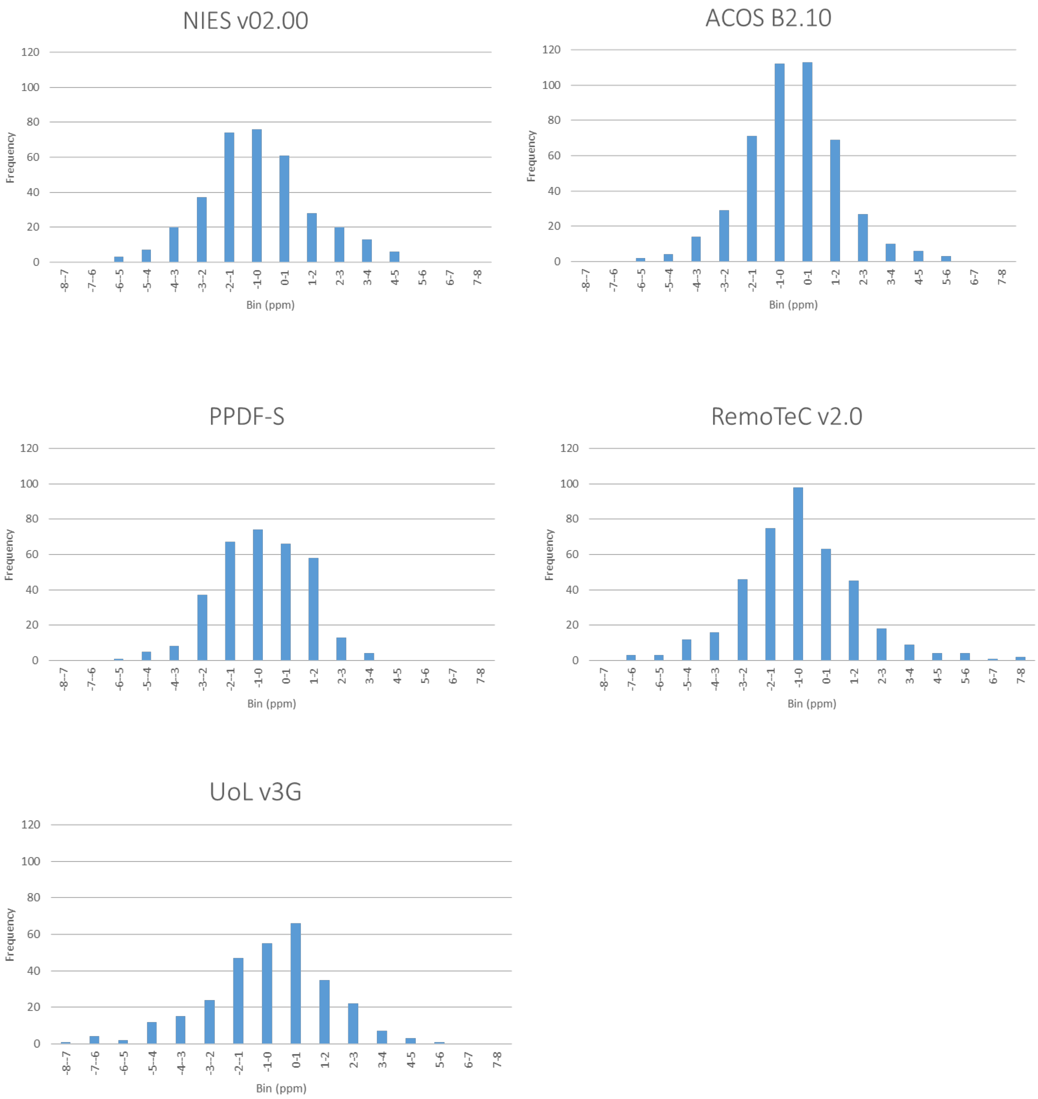

Figure 4.4. Frequency distribution of GOSAT-TCCON differences in ppm (NIES v02.00, ACOS B2.10, PPDF-S, RemoTeC v2.0, and UoL v3G). Abscissa: concentration bins in ppm (bin size: 1 ppm).

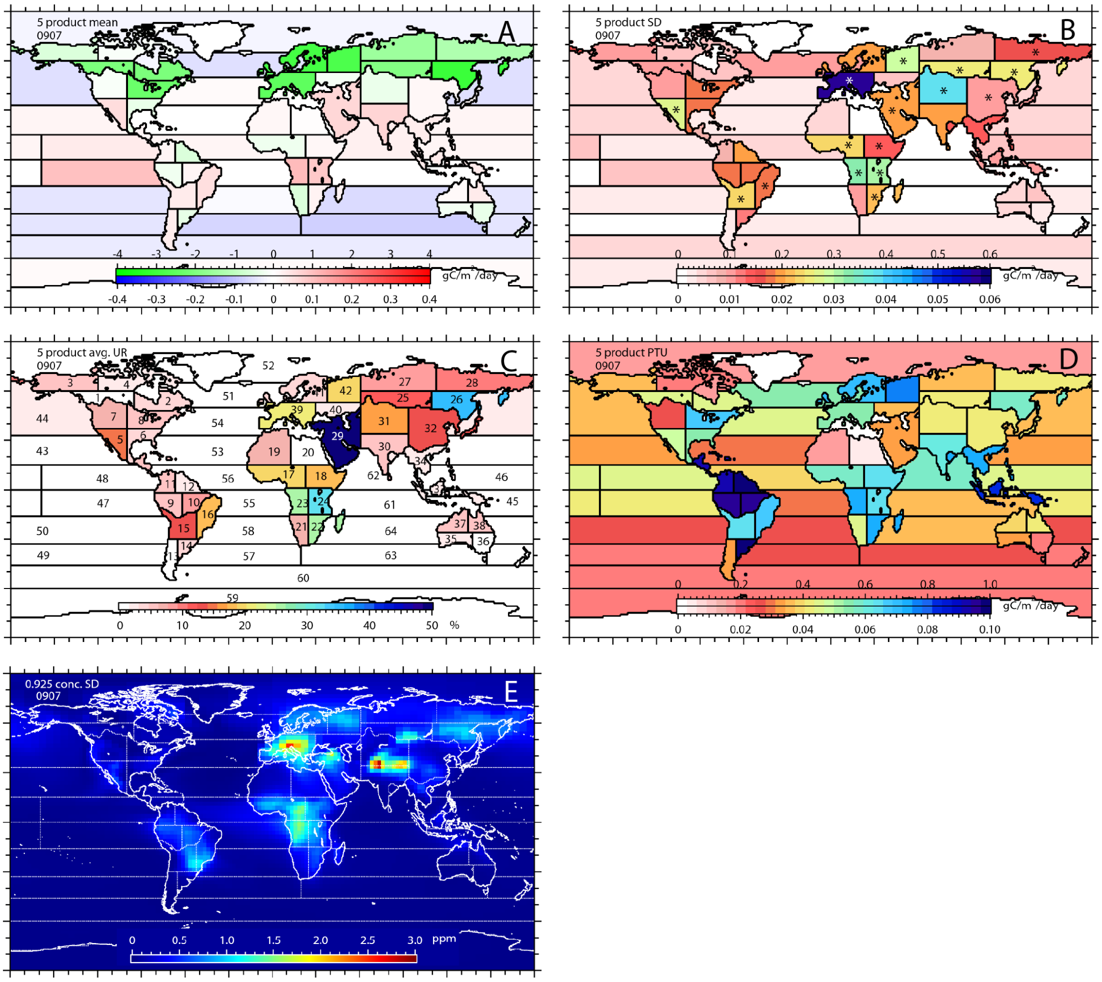

Figure 4.5. Panels A and B: mean and standard deviation of five independent monthly flux estimates for July 2009 (in gC m^{-2} day^{-1}). Panel C: mean of five uncertainty reduction rates (UR; %) relative to GV-only inversion. The printed value in each region represents region ID number, and the color denotes uncertainty reduction magnitude. Asterisks in panel B indicates regions with UR ≥10%. Panel D: Mean of five a posteriori uncertainties. Panel E: SD of five a posteriori concentrations (in ppm; monthly-mean concentrations simulated on 2.5°×2.5° grid at 0.975 sigma level). The upper and lower scales embedded in panels A and B are for the terrestrial and oceanic values, respectively. Note the oceanic scale is one tenth of the terrestrial one.

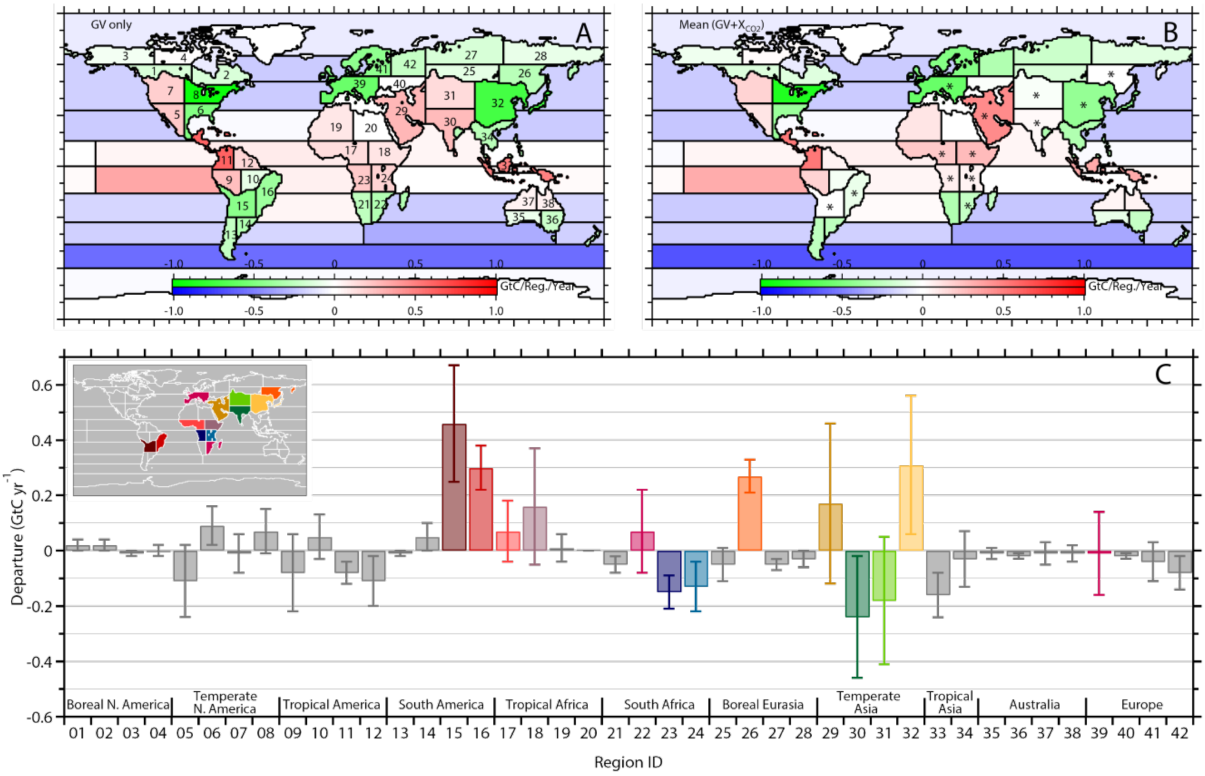

Figure 4.6. Panels A and B: annual mean regional fluxes estimated from GV data alone and both GV and GOSAT X_{CO2} retrievals, respectively (in GtC region^{-1} yr^{-1}). Anthropogenic and biomass burning emissions are not included. Panel C: mean of the departure of five GV+X_{CO2} estimates from the GV-only results (in GtC region^{-1} yr^{-1}). Colored values are associated with the pronounced influence of GOSAT retrievals (mean URs ≥ 10%). Error bar: SD of five departure values. Inset on panel C indicates the locations of the high-UR regions.

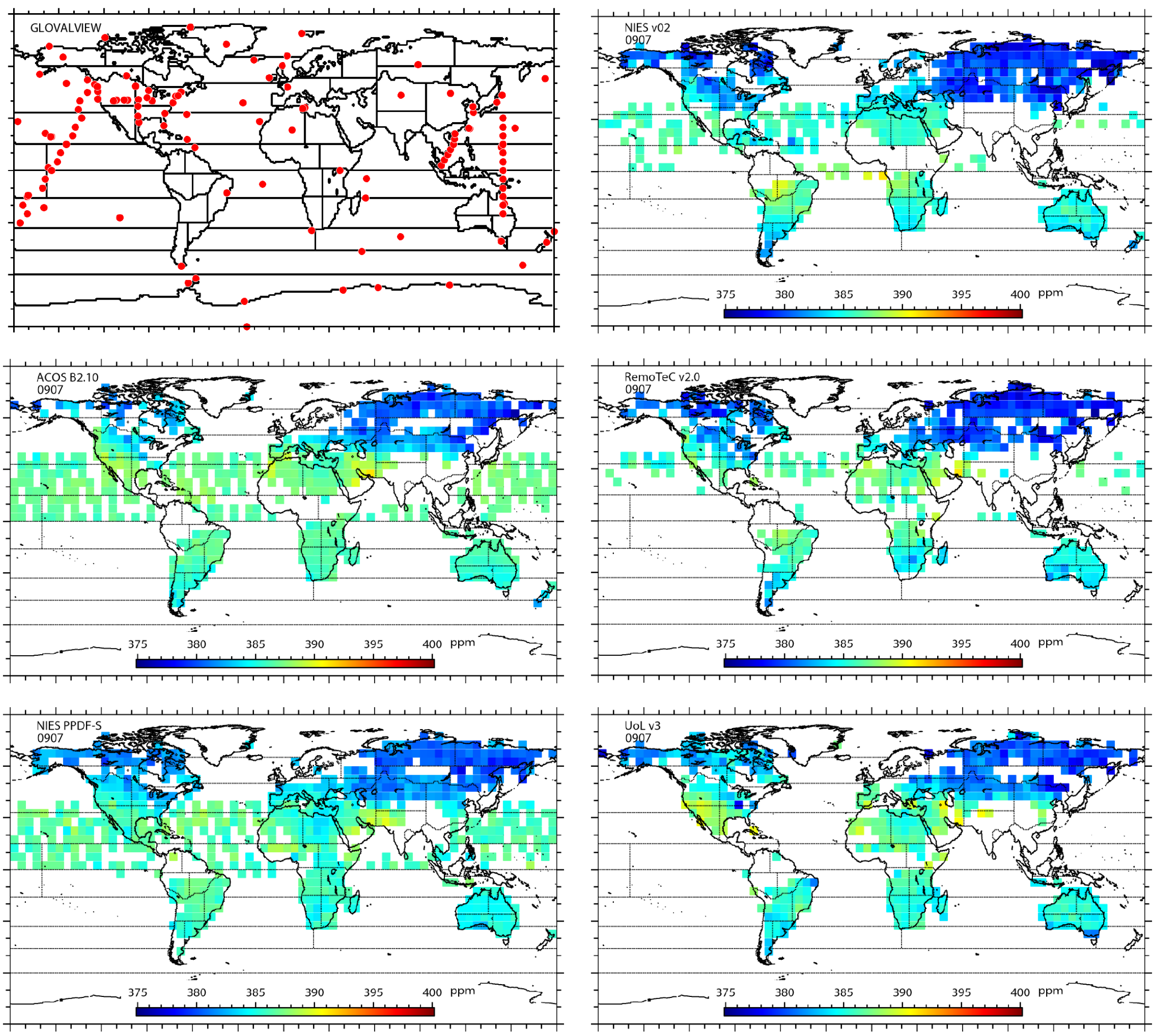

Figure 4.7. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms. Values shown: July 2009.

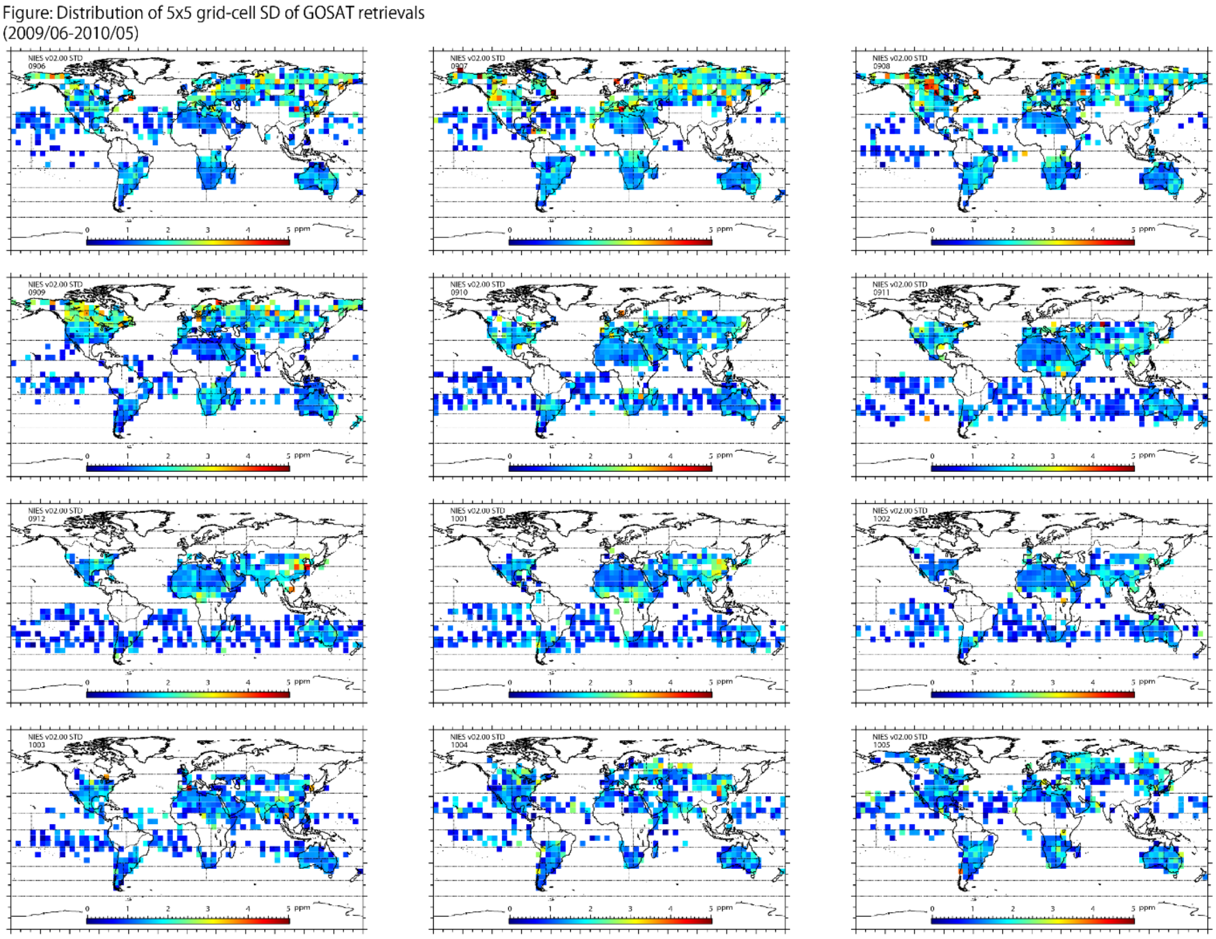

Figure 4S.1. SDs of X_{CO2} retrievals found in 5°×5° grid cells (NIES v02.00) over the 12-months analysis period.

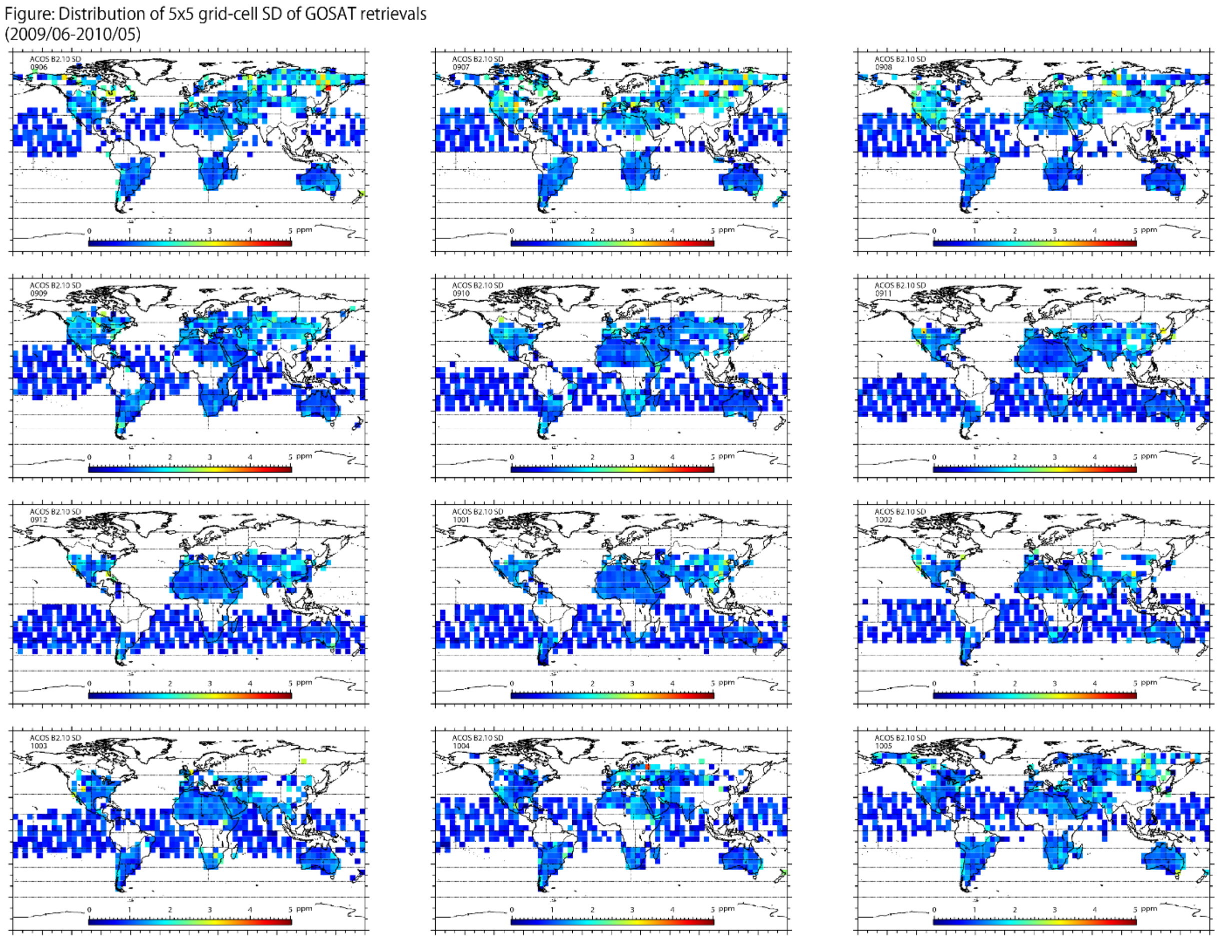

Figure 4S.2. SDs of X_{CO2} retrievals found in 5°×5° grid cells (ACOS B2.10) over the 12- months analysis period.

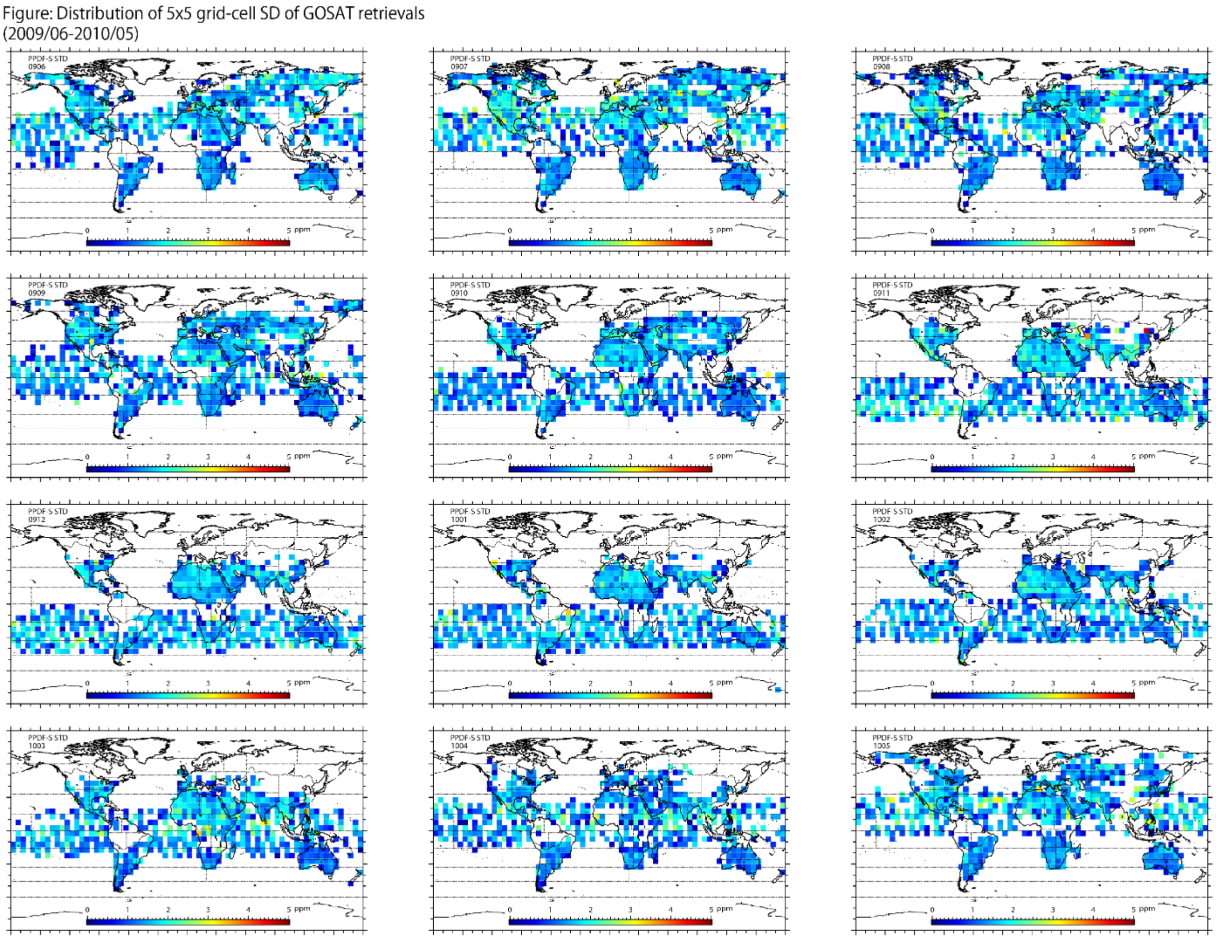

Figure 4S.3. SDs of X_{CO2} retrievals found in 5°×5° grid cells (PPDF-S) over the 12-months analysis period.

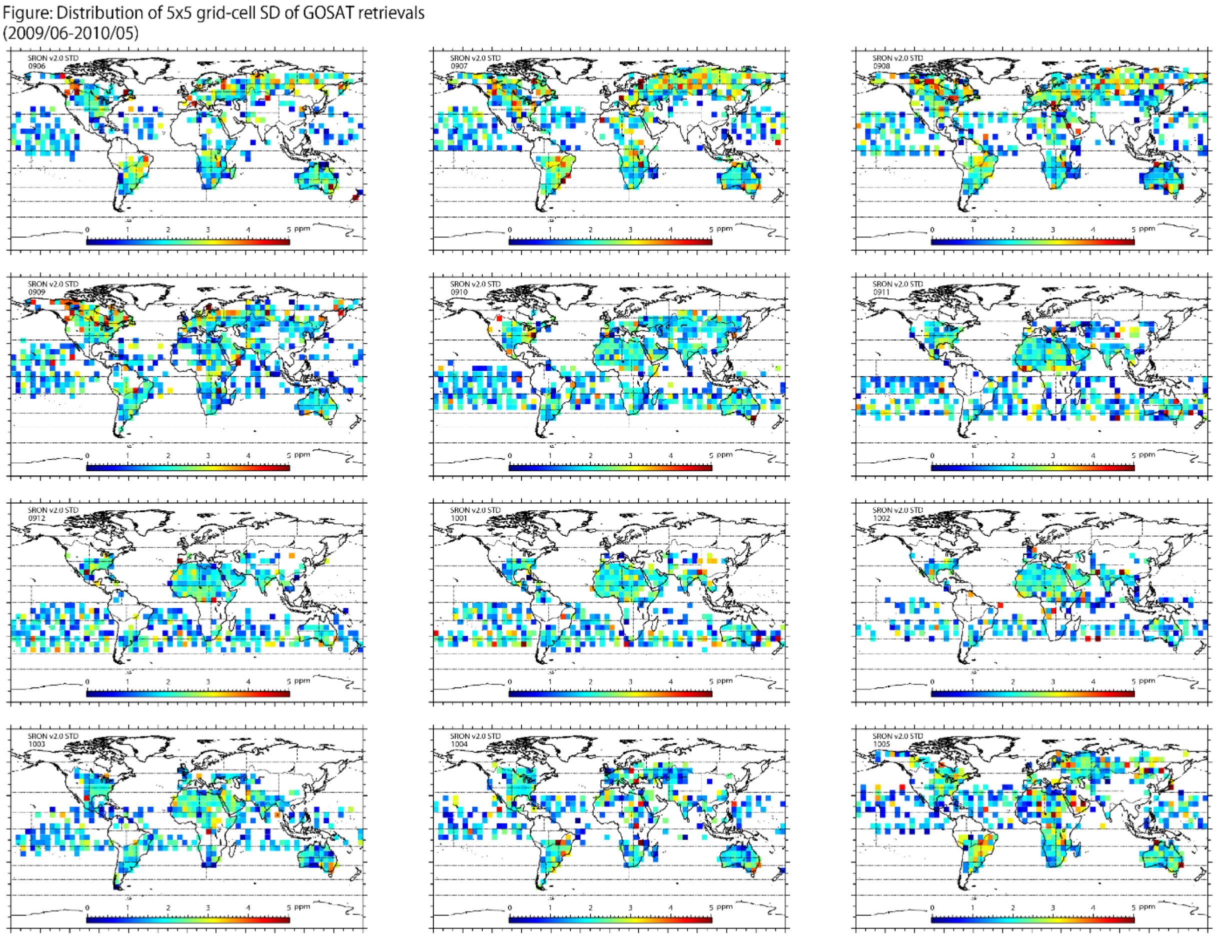

Figure 4S.4. SDs of X_{CO2} retrievals found in 5°×5° grid cells (RemoTeC v2.0) over the 12- months analysis period.

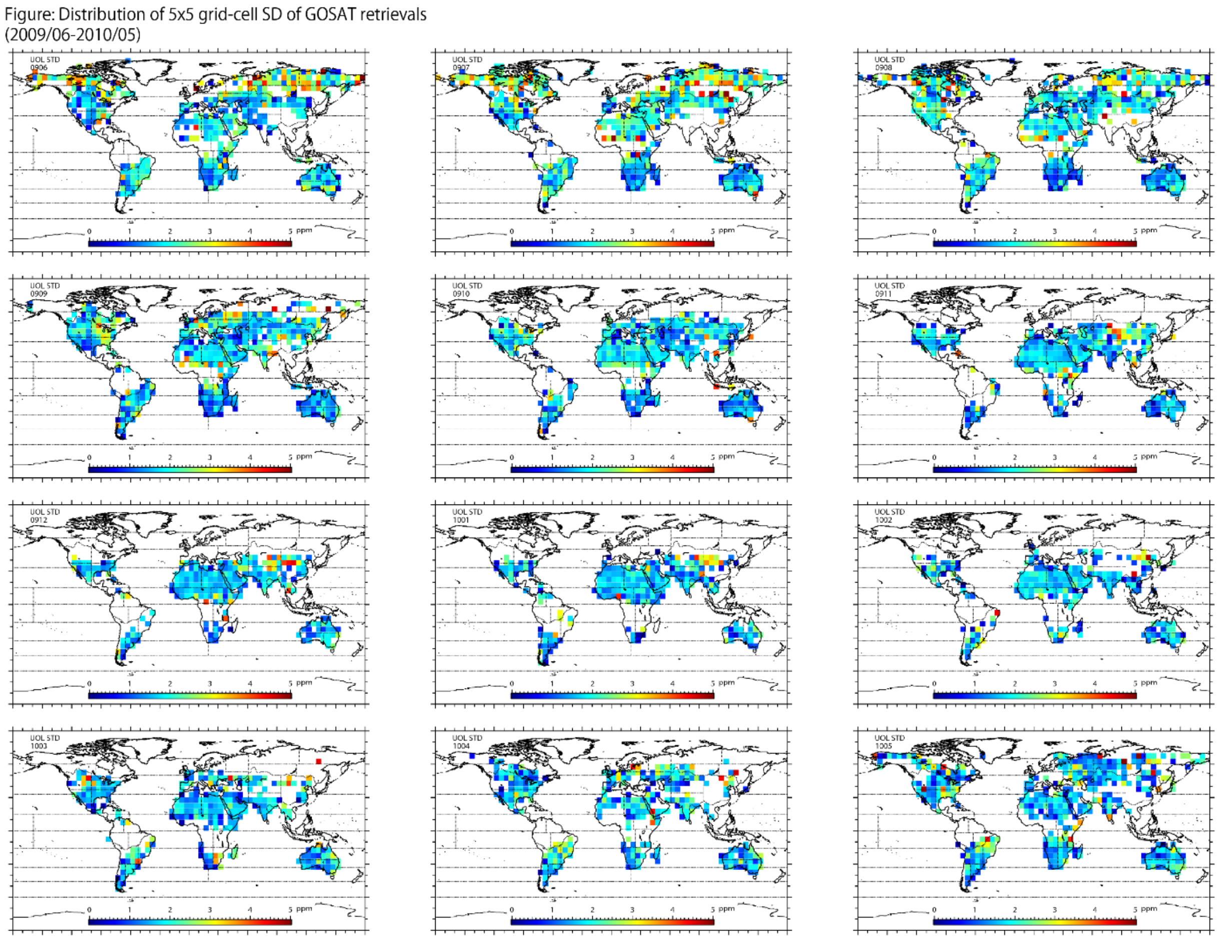

Figure 4S.4. SDs of X_{CO2} retrievals found in 5°×5° grid cells (UoL v3G) over the 12-months analysis period.

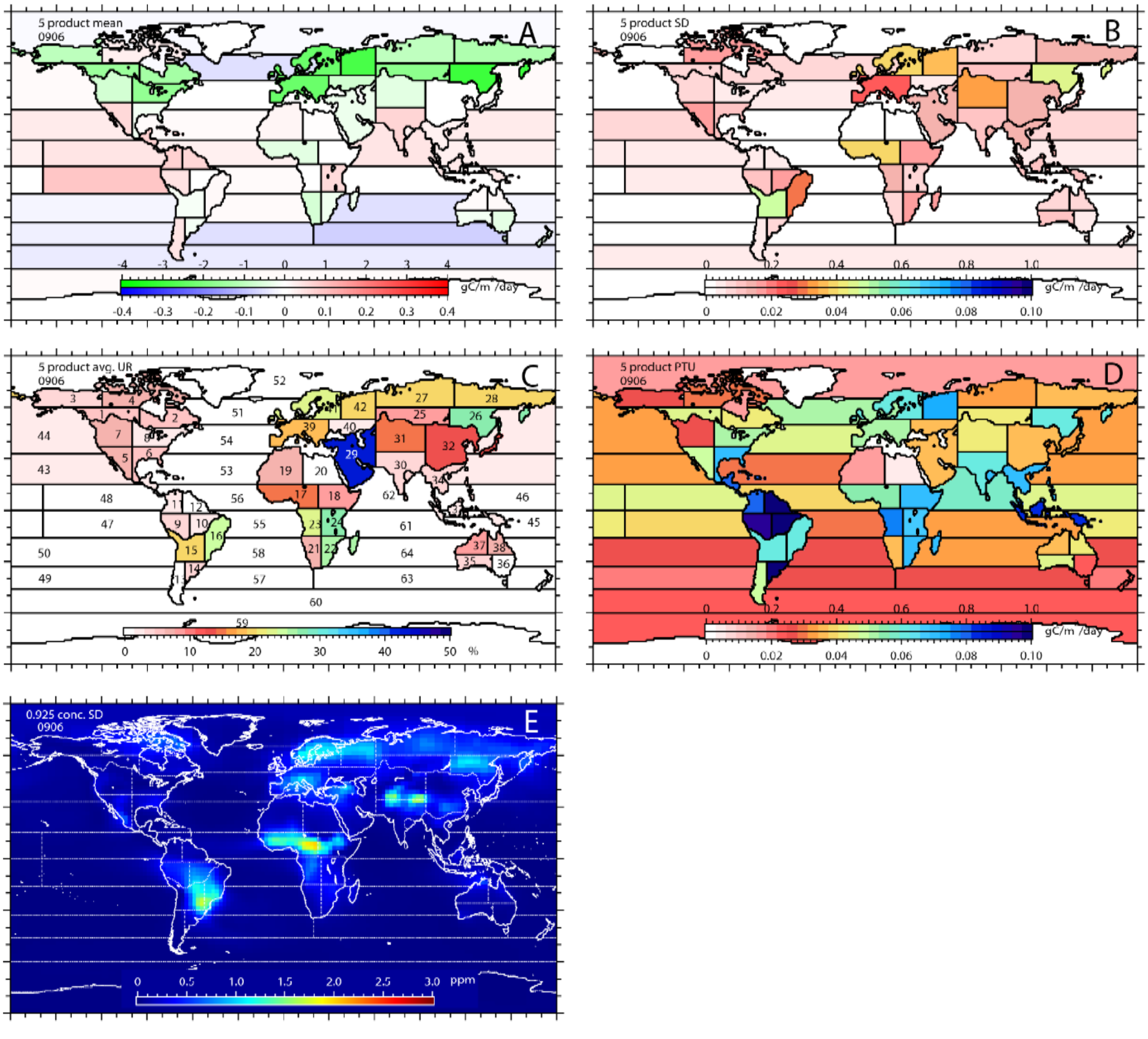

Figure 4S.5. Figure 4.5 for June 2009. See caption for Figure 4.5 for explanation.

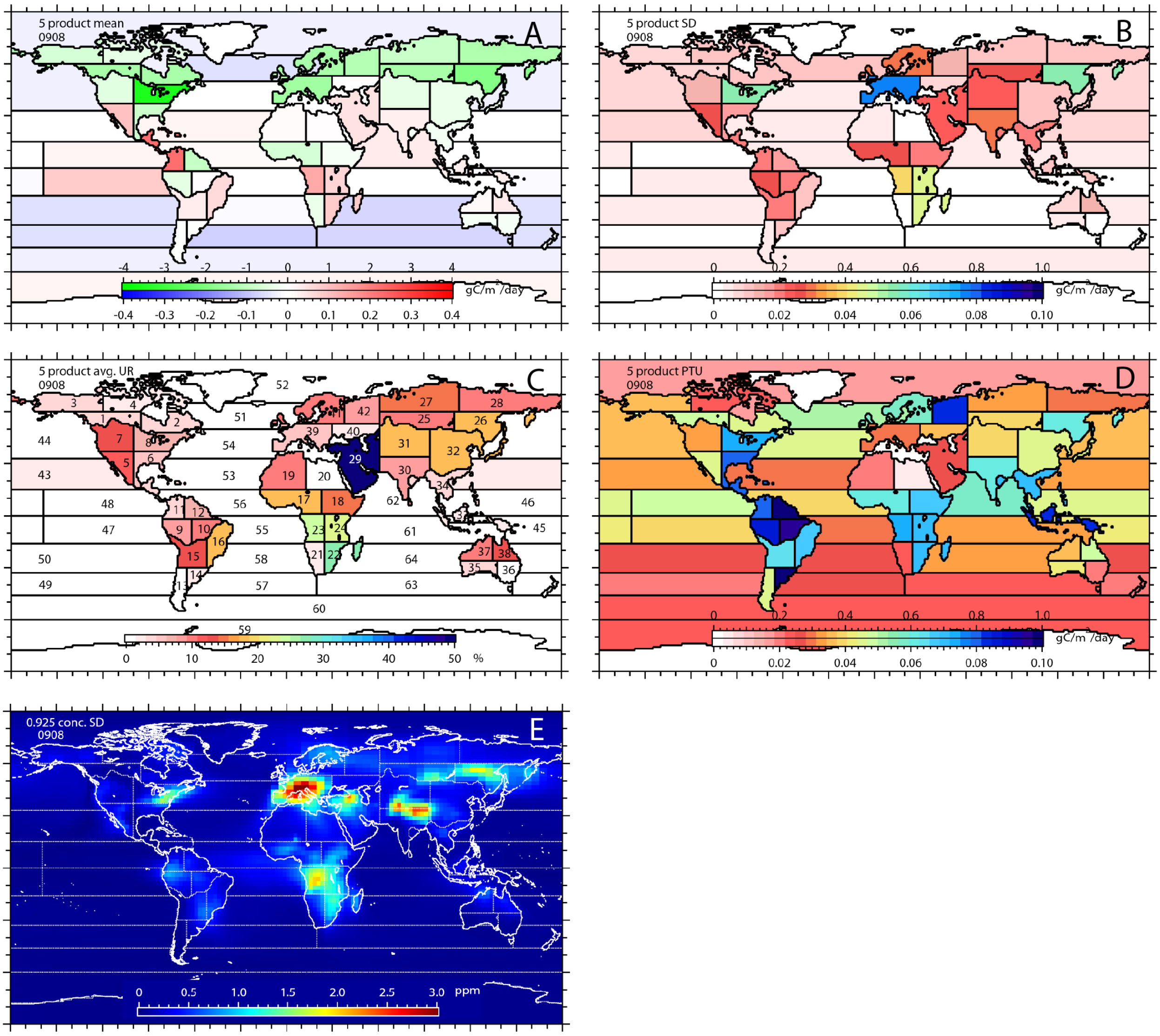

Figure 4S.6. Figure 4.5 for August 2009. See caption for Figure 4.5 for explanation.

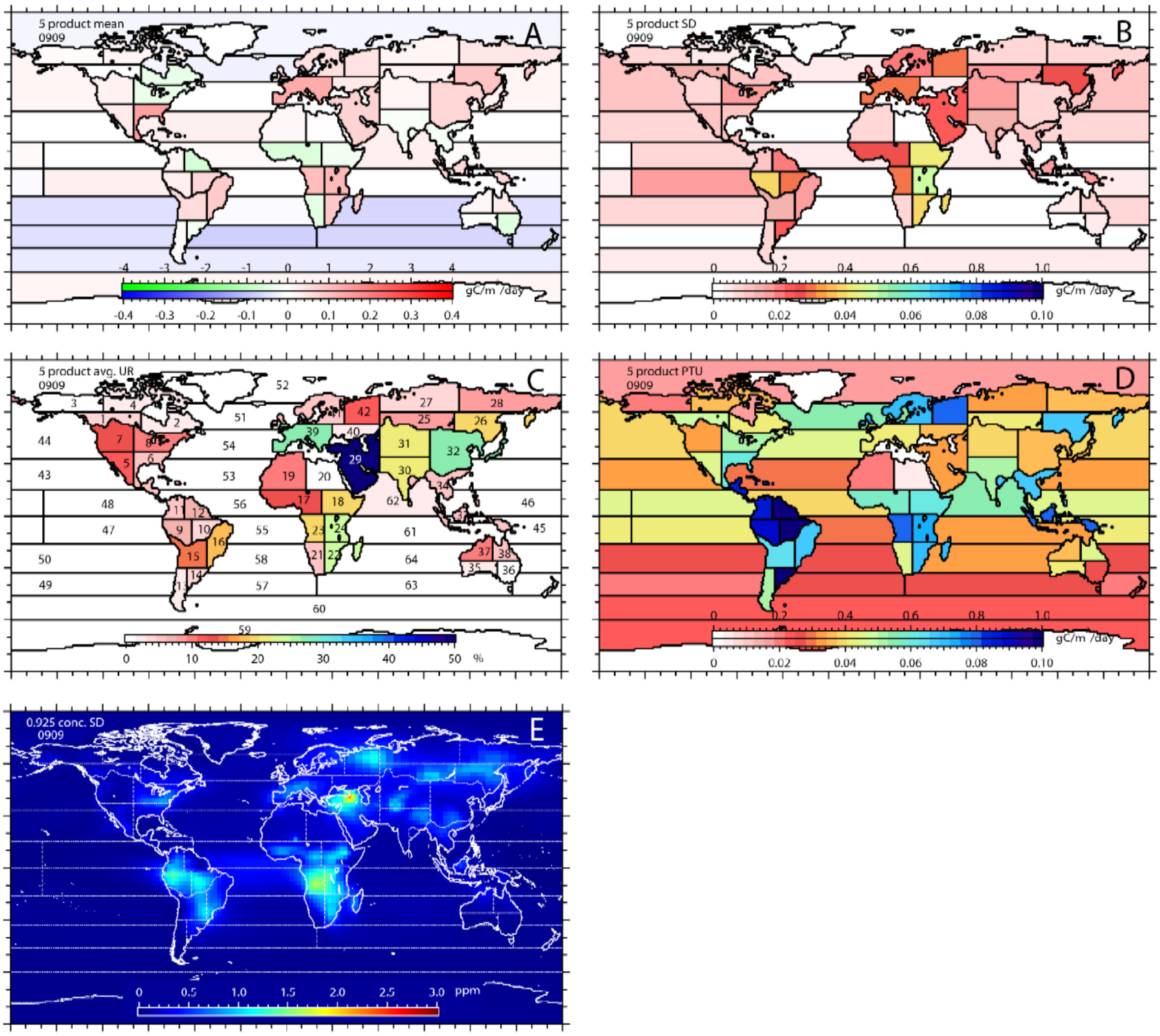

Figure 4S.7. Figure 4.5 for September 2009. See caption for Figure 4.5 for explanation.

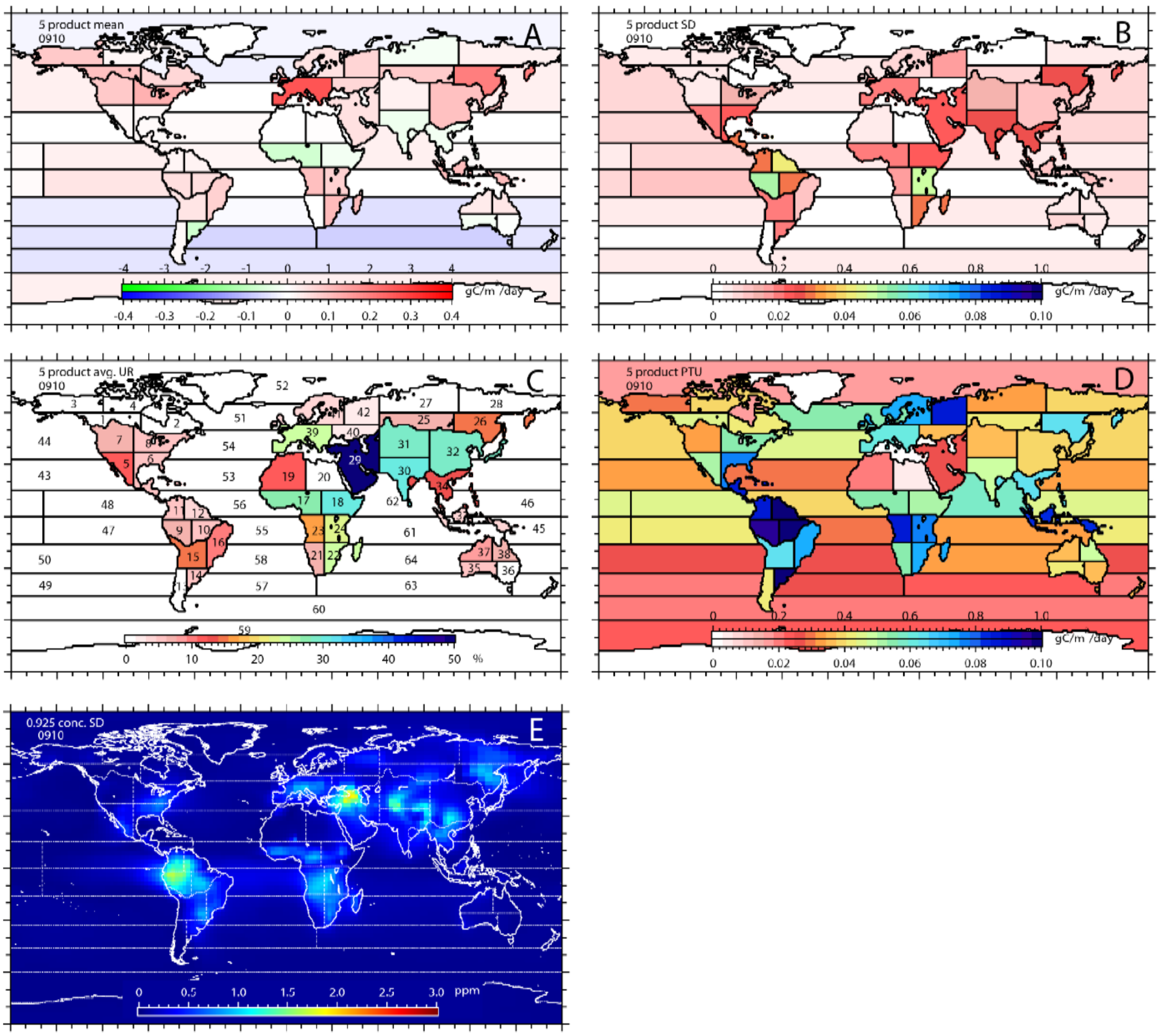

Figure 4S.8. Figure 4.5 for October 2009. See caption for Figure 4.5 for explanation.

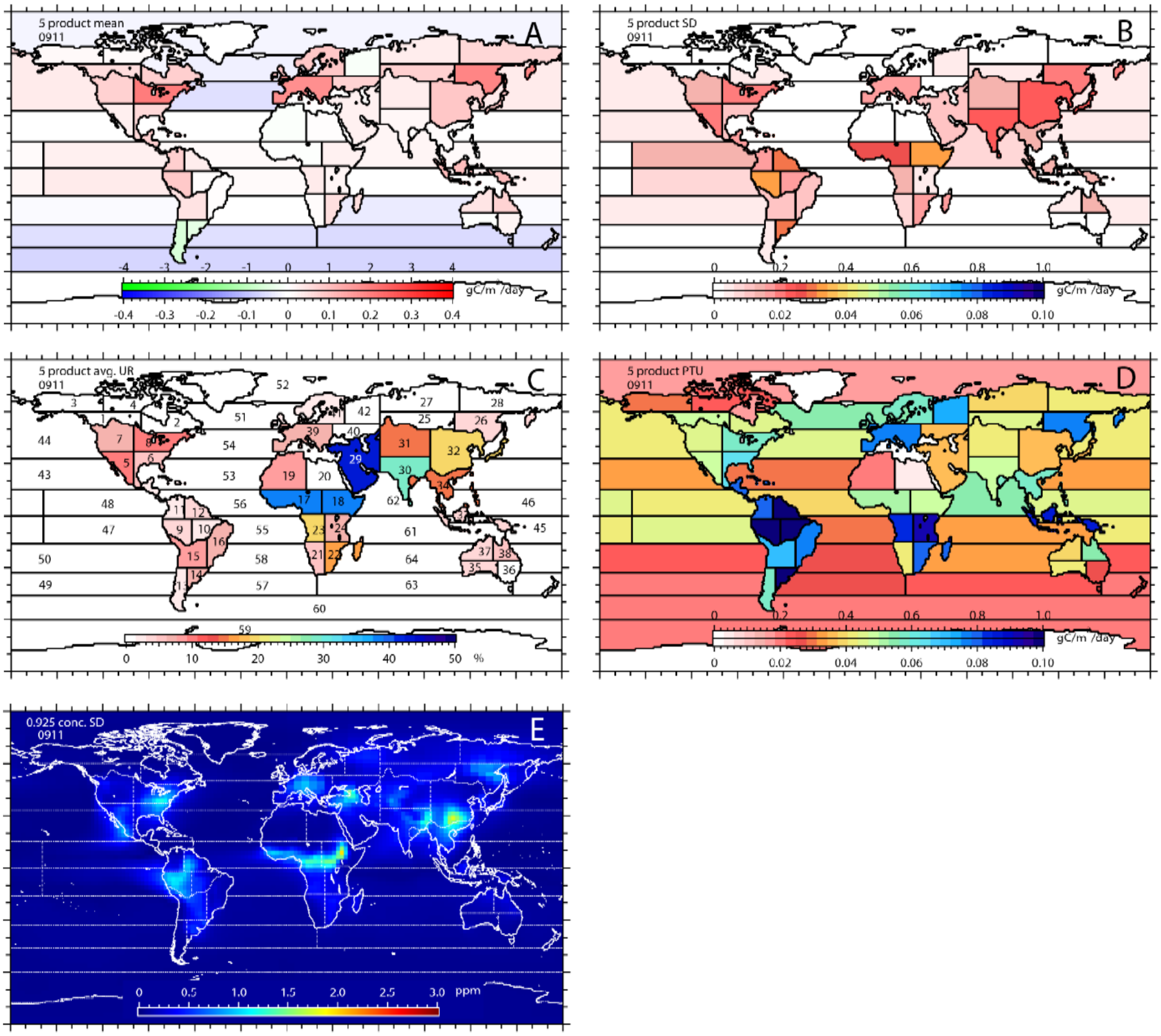

Figure 4S.9. Figure 4.5 for November 2009. See caption for Figure 4.5 for explanation.

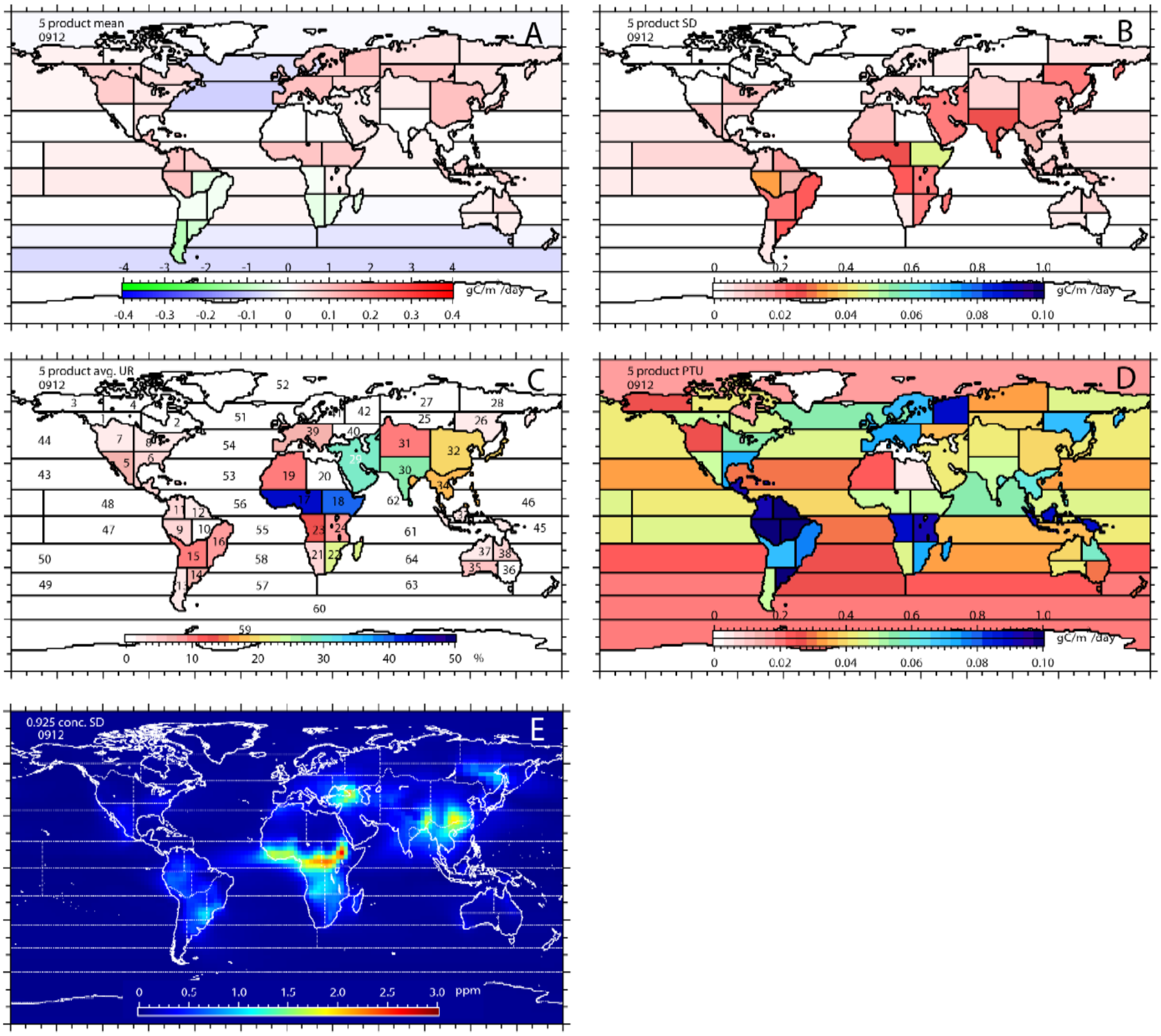

Figure 4S.10. Figure 4.5 for December 2009. See caption for Figure 4.5 for explanation.

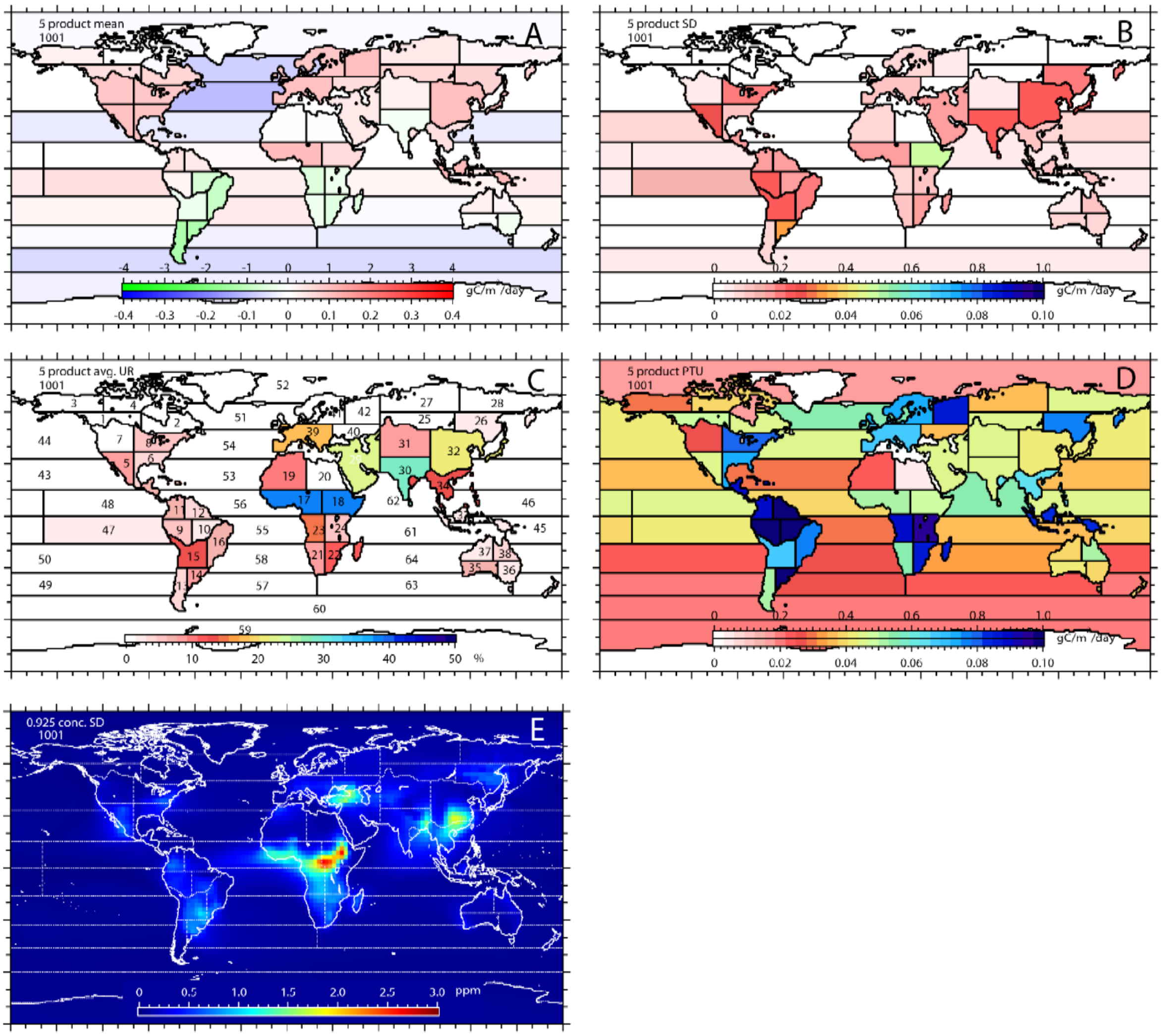

Figure 4S.11. Figure 4.5 for January 2010. See caption for Figure 4.5 for explanation.

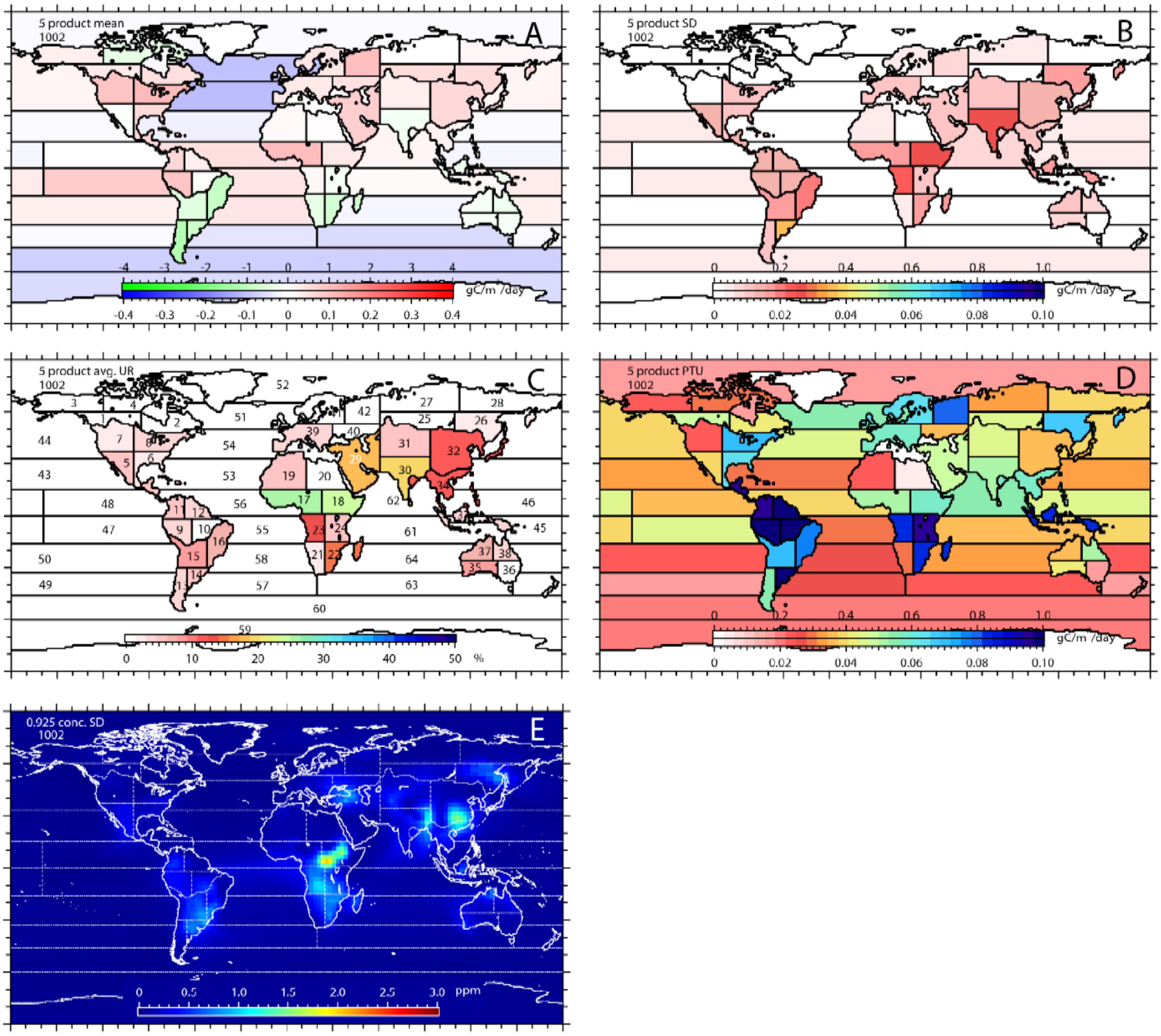

Figure 4S.12. Figure 4.5 for February 2010. See caption for Figure 4.5 for explanation.

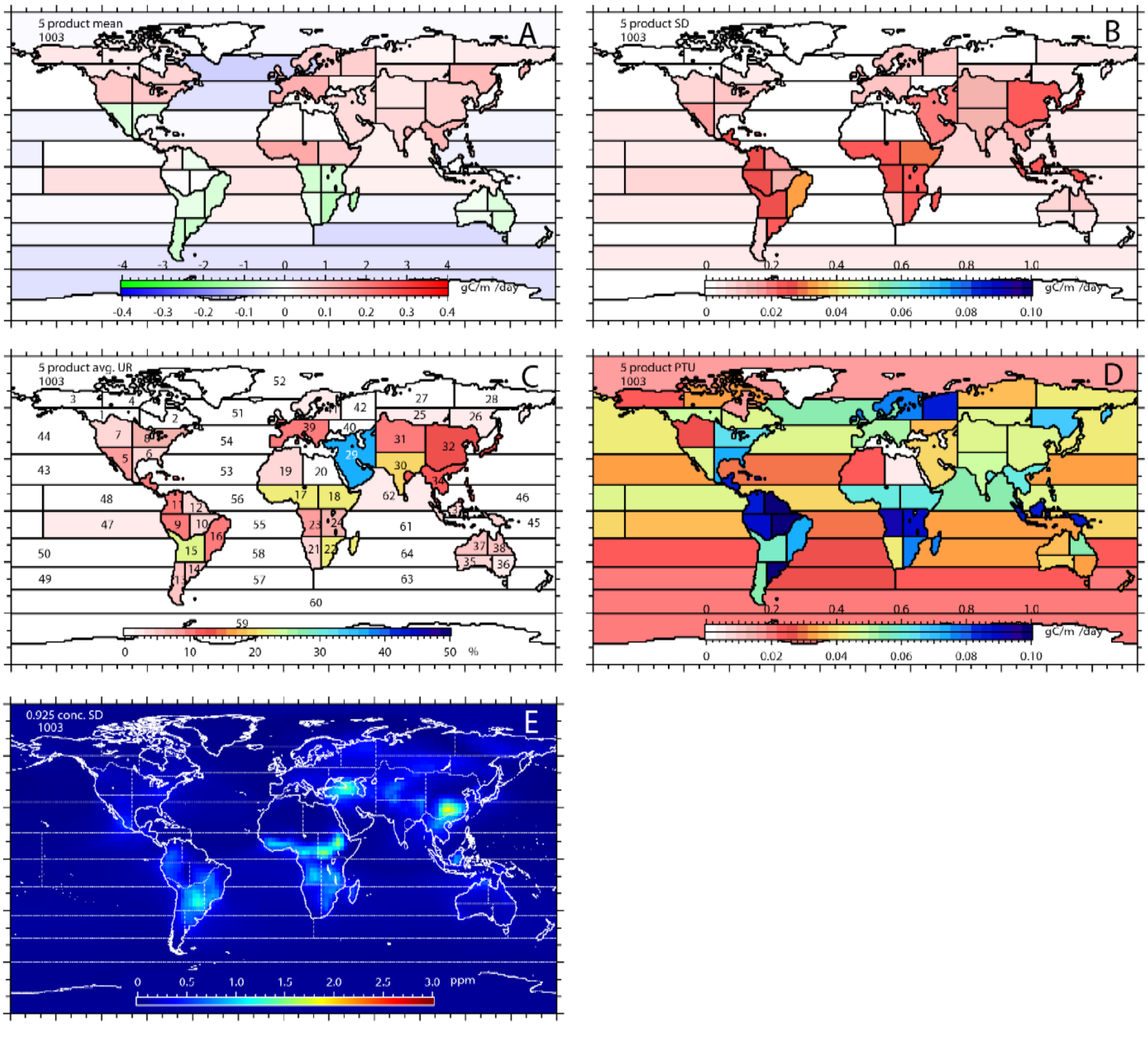

Figure 4S.13. Figure 4.5 for March 2010. See caption for Figure 4.5 for explanation.

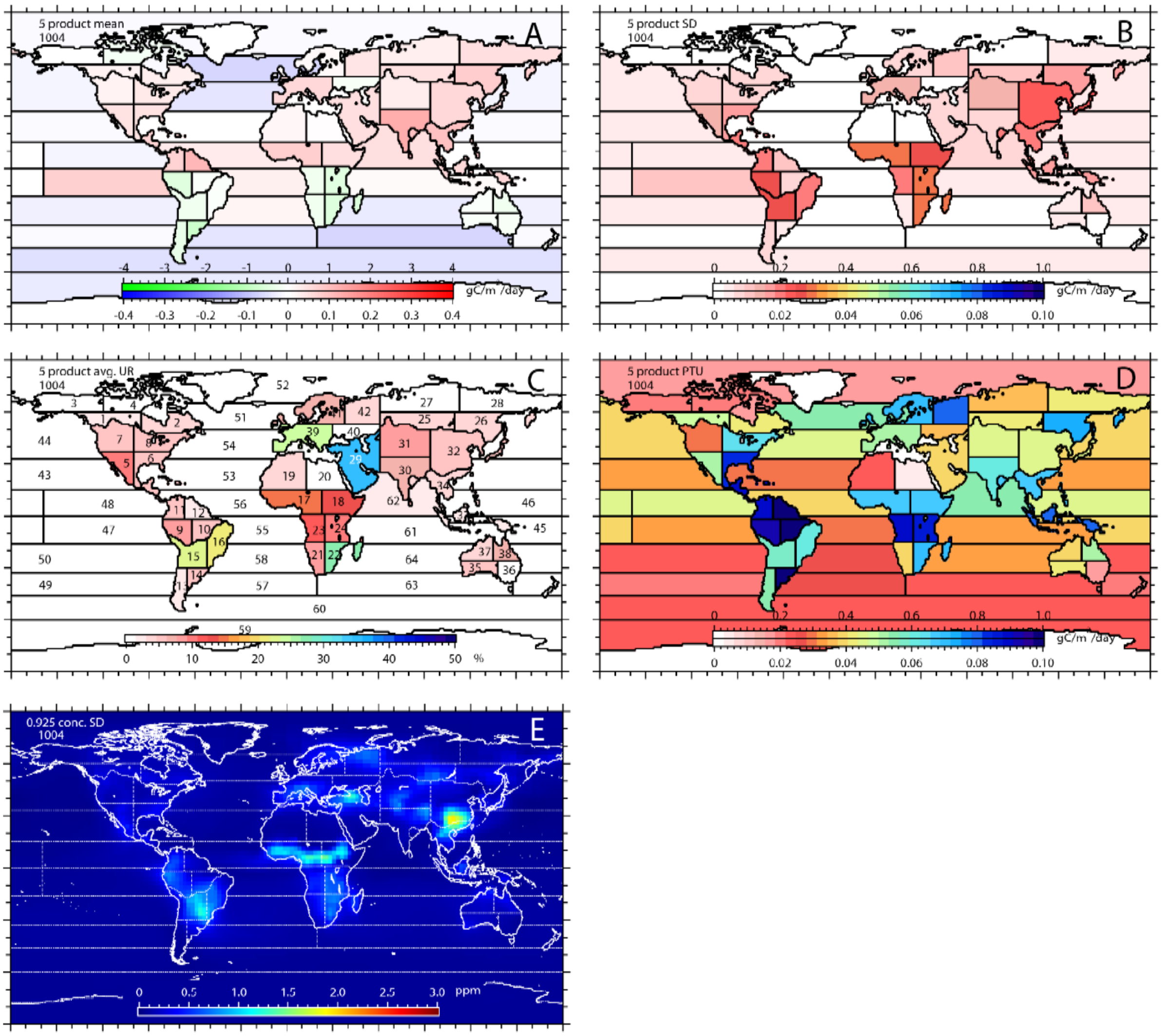

Figure 4S.14. Figure 4.5 for April 2010. See caption for Figure 4.5 for explanation.

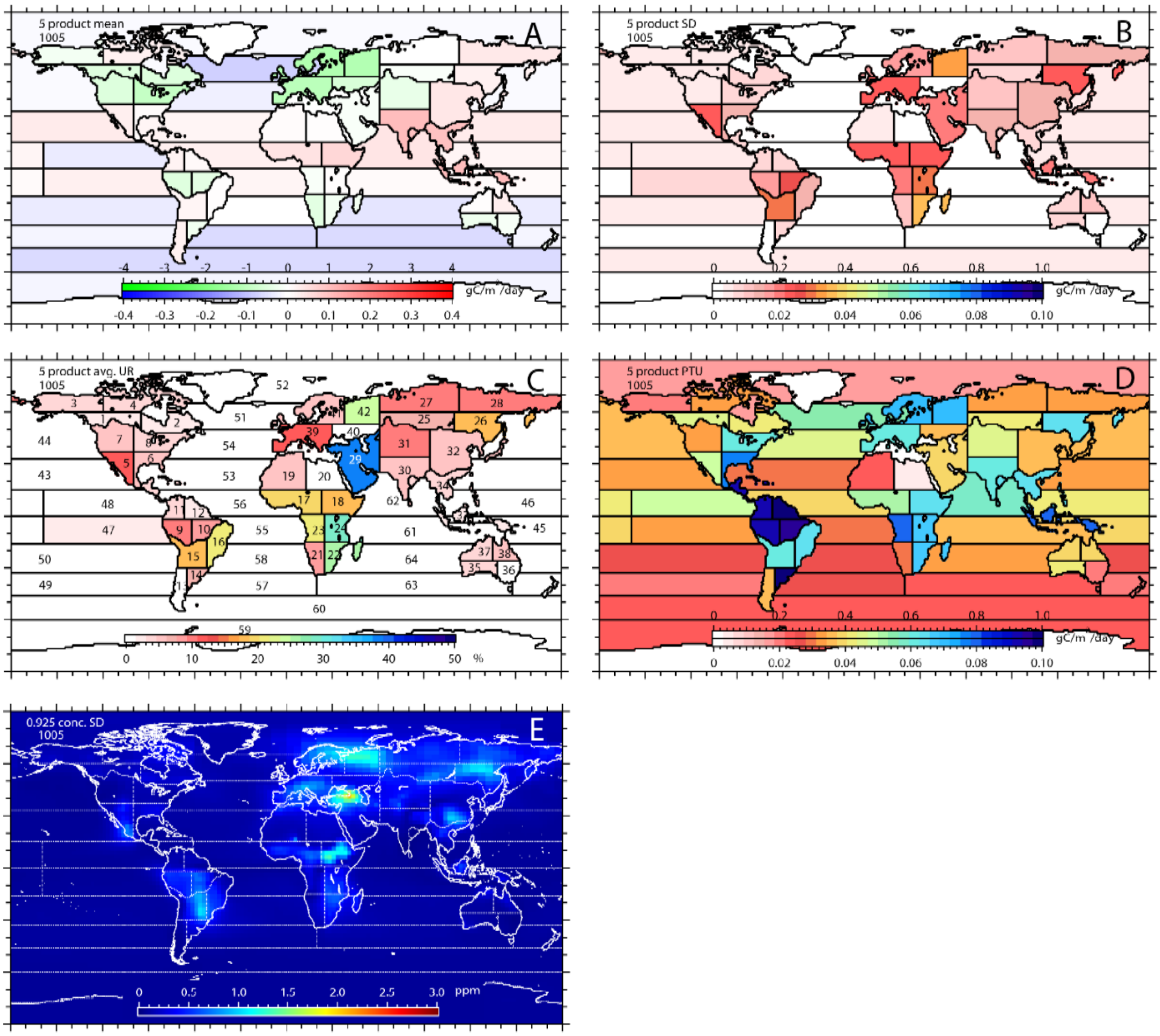

Figure 4S.15. Figure 4.5 for May 2010. See caption for Figure 4.5 for explanation.

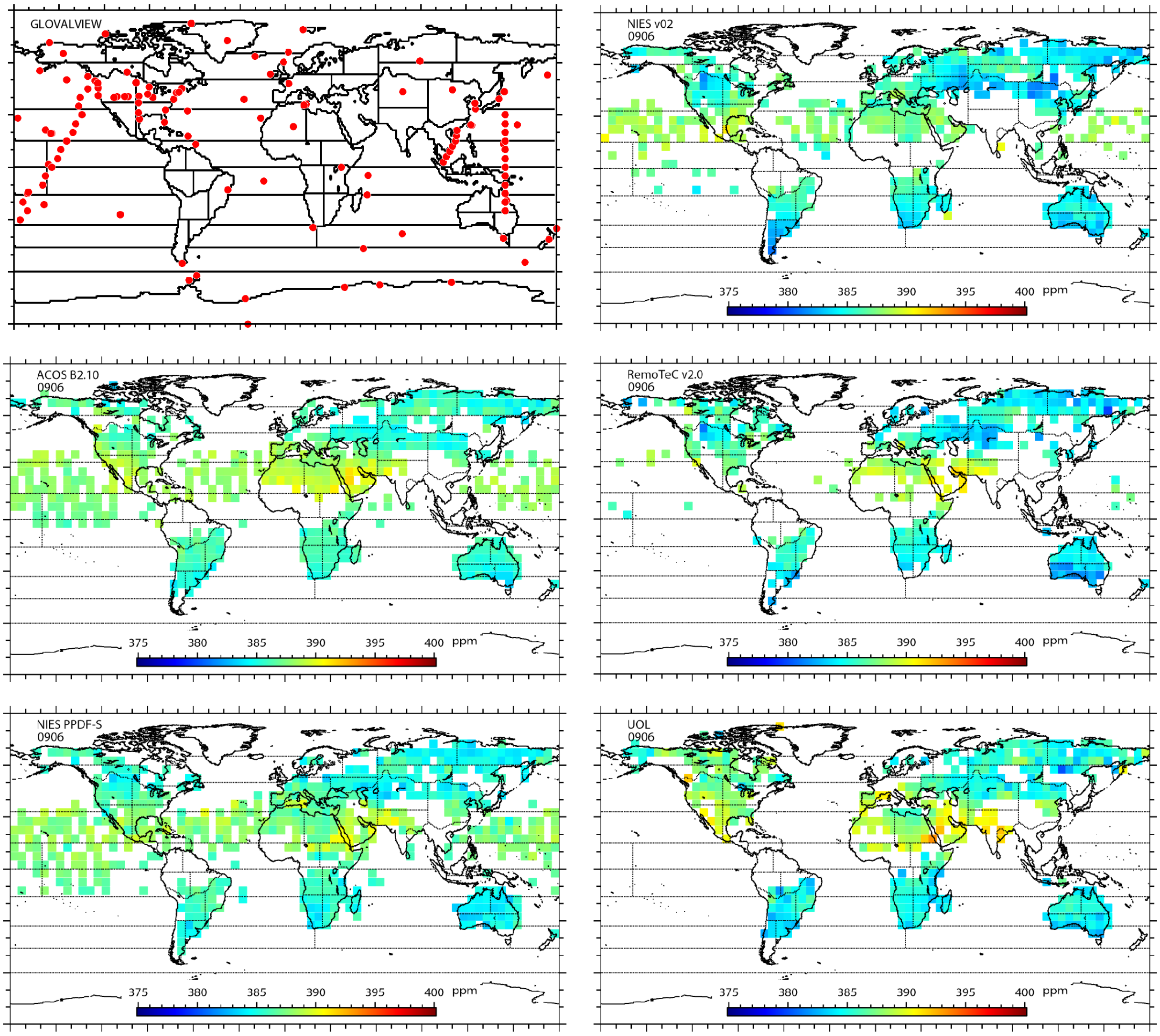

Figure 4S.16. Figure 4.7 for June 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

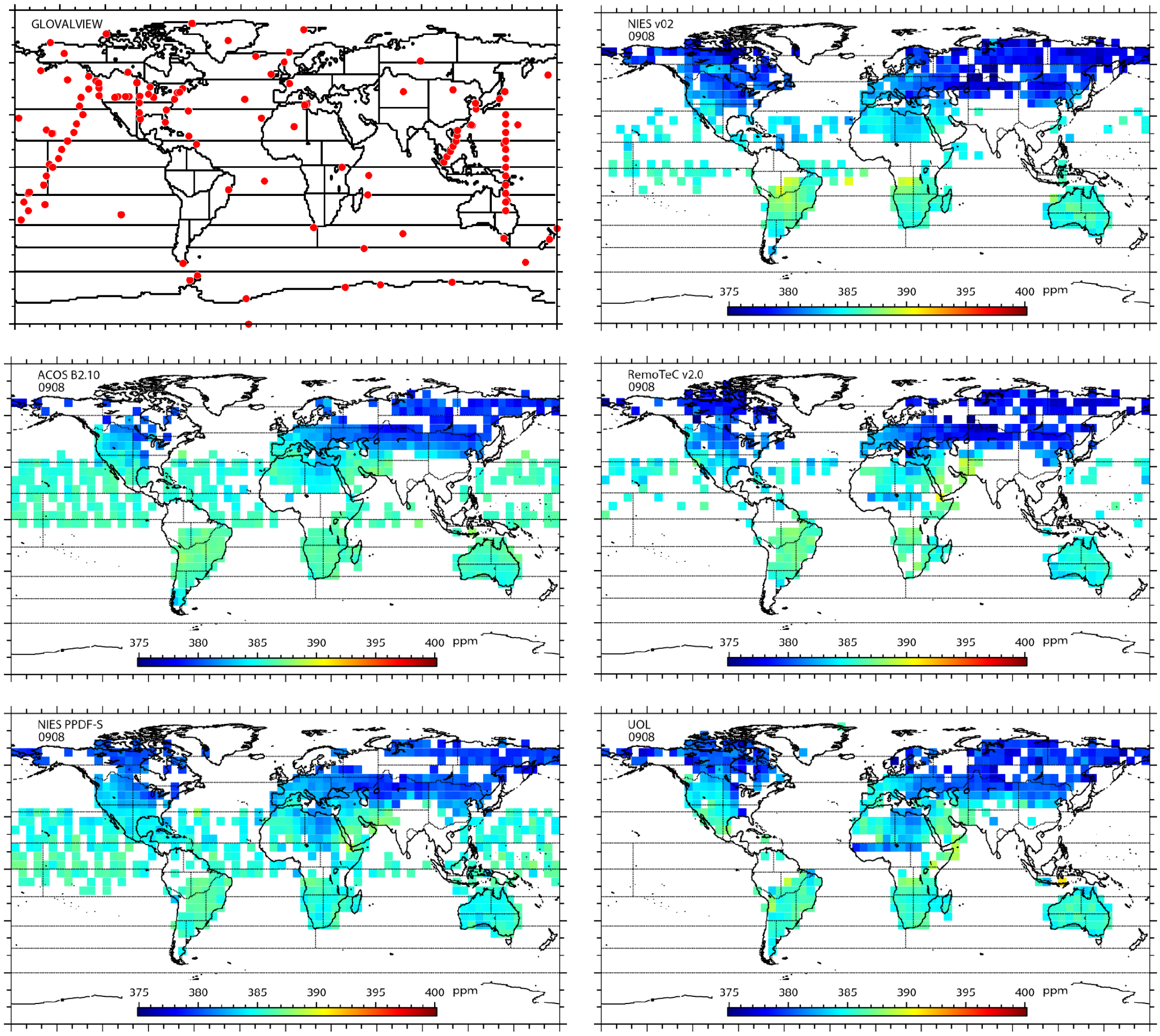

Figure 4S.17. Figure 4.7 for August 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

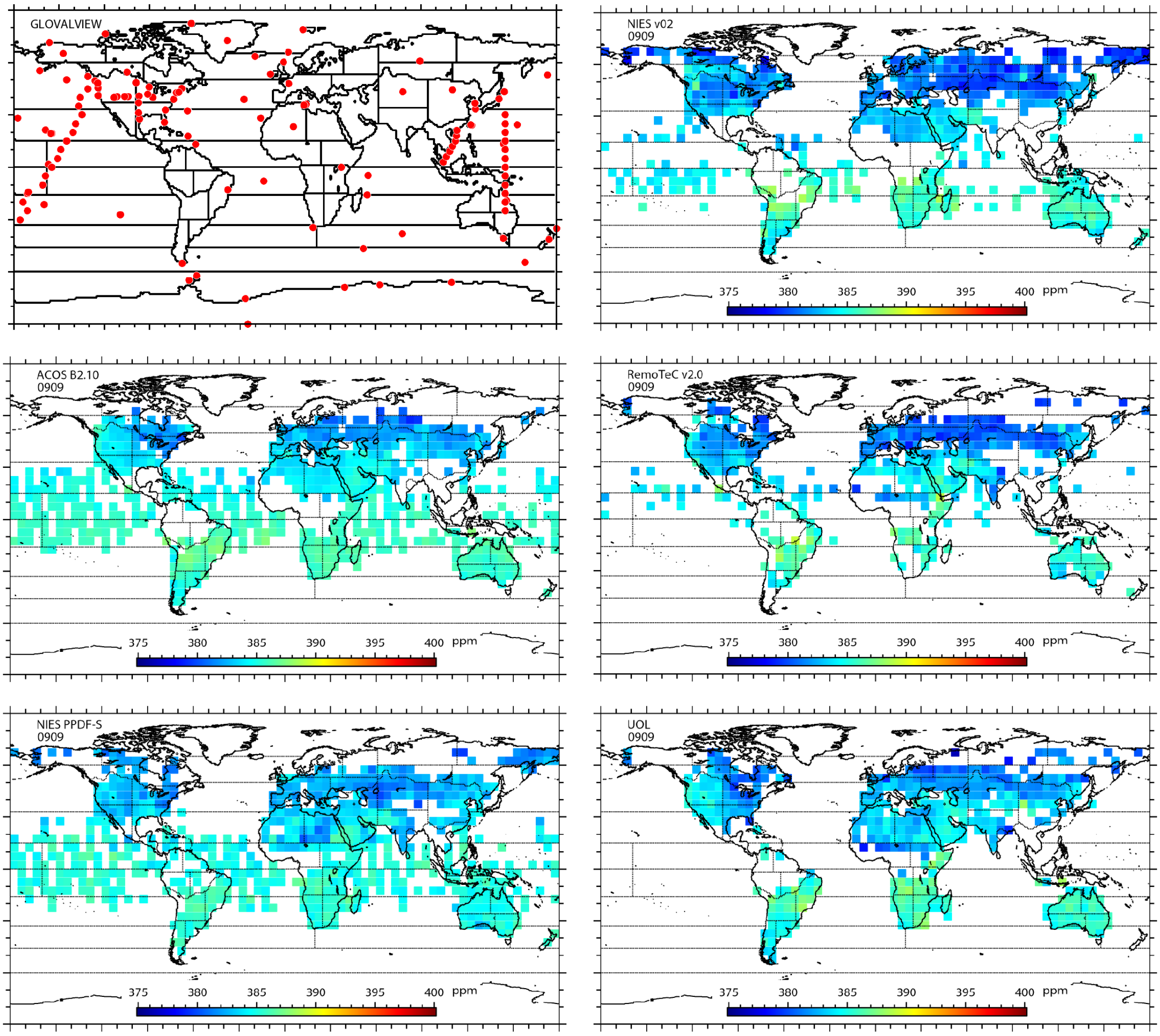

Figure 4S.18. Figure 4.7 for September 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

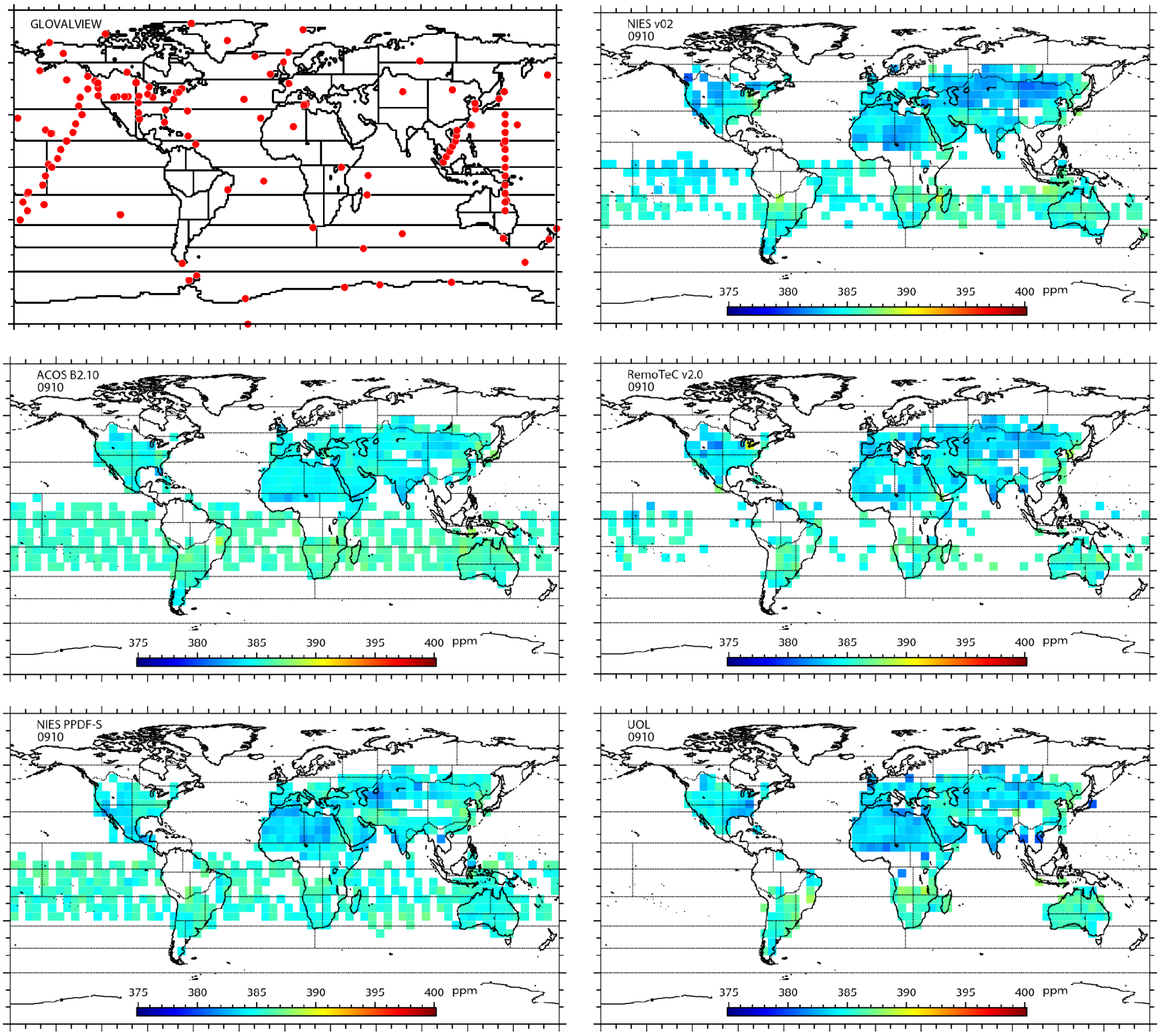

Figure 4S.19. Figure 4.7 for October 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

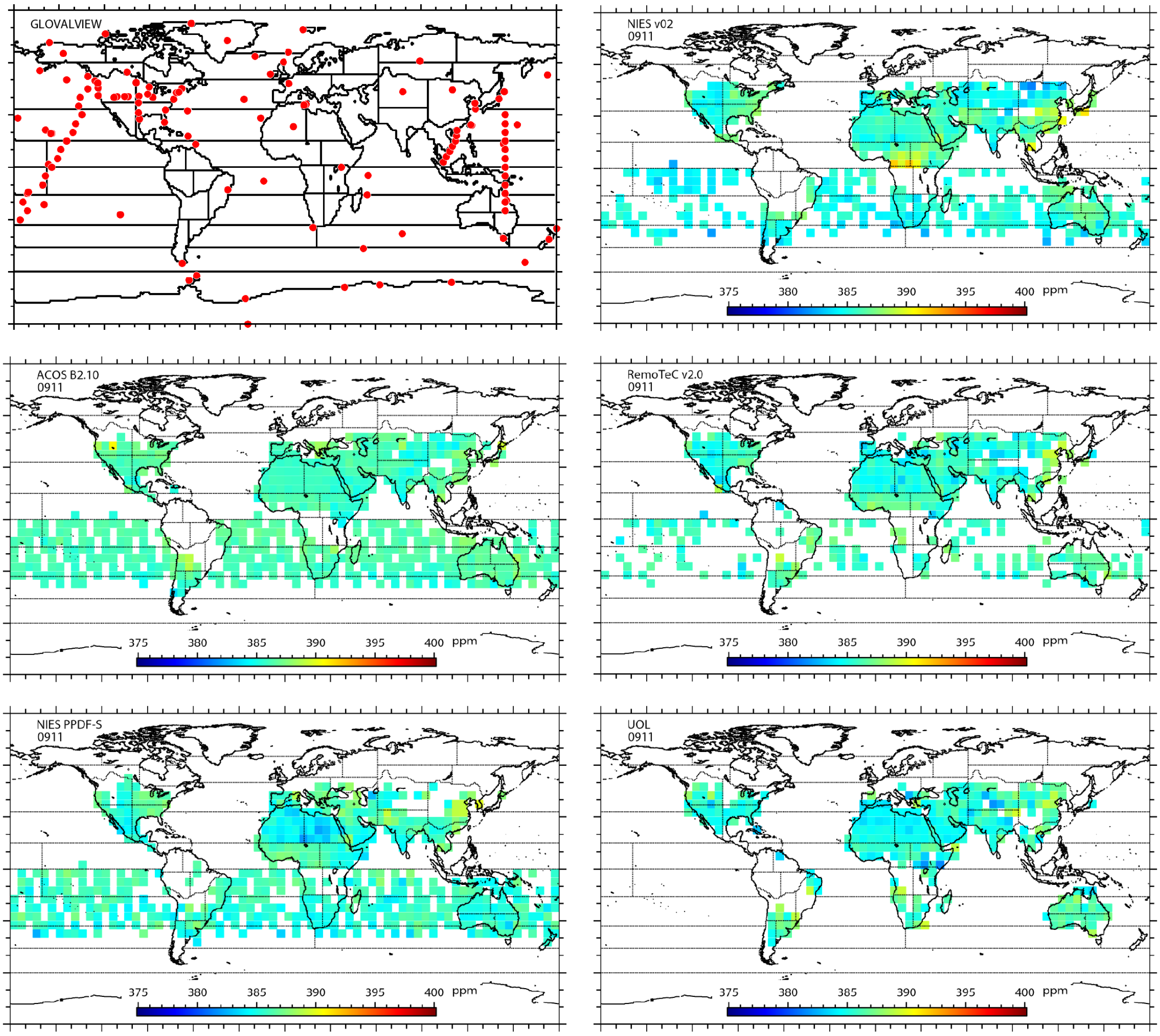

Figure 4S.20. Figure 4.7 for November 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

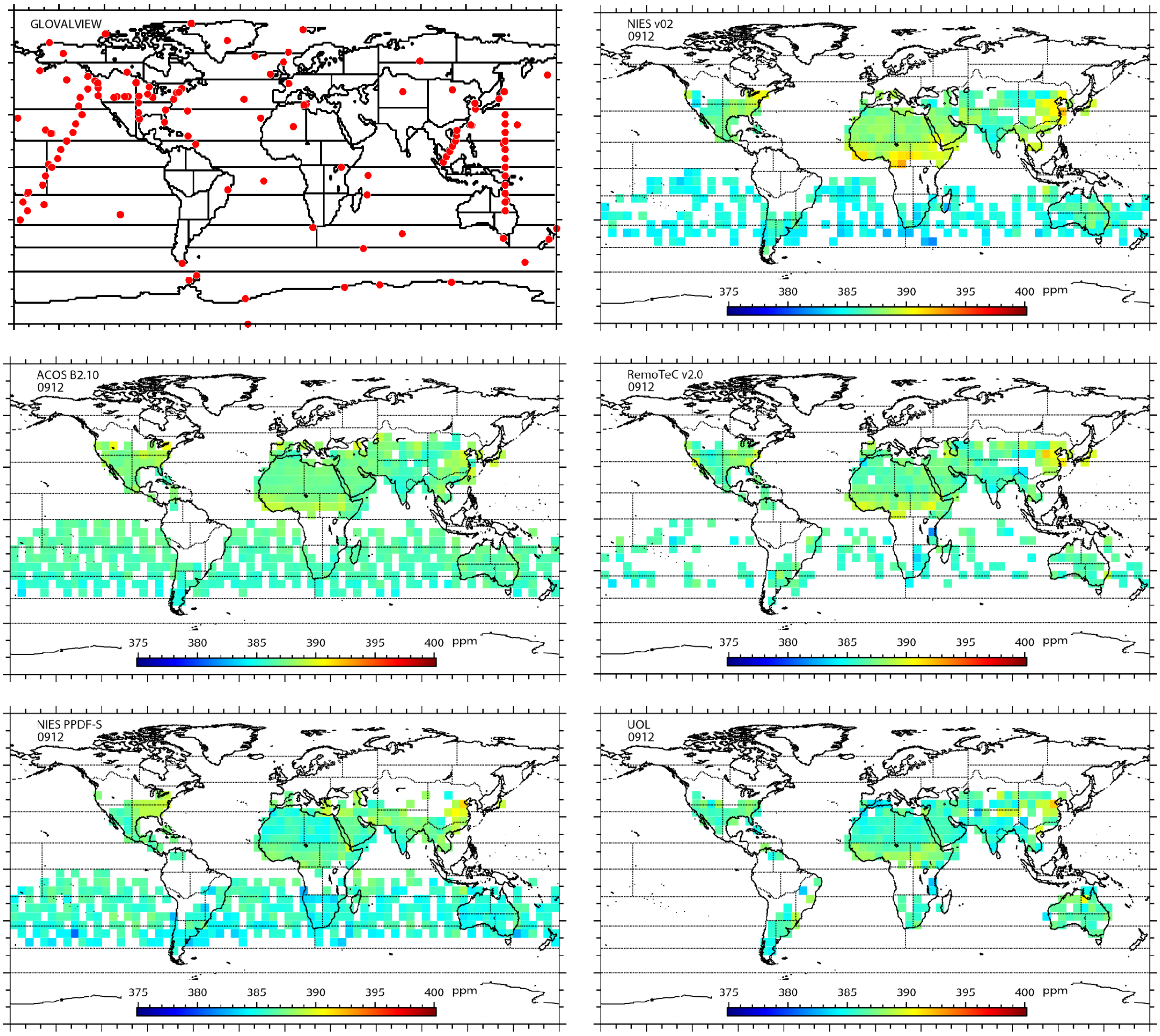

Figure 4S.21. Figure 4.7 for December 2009. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

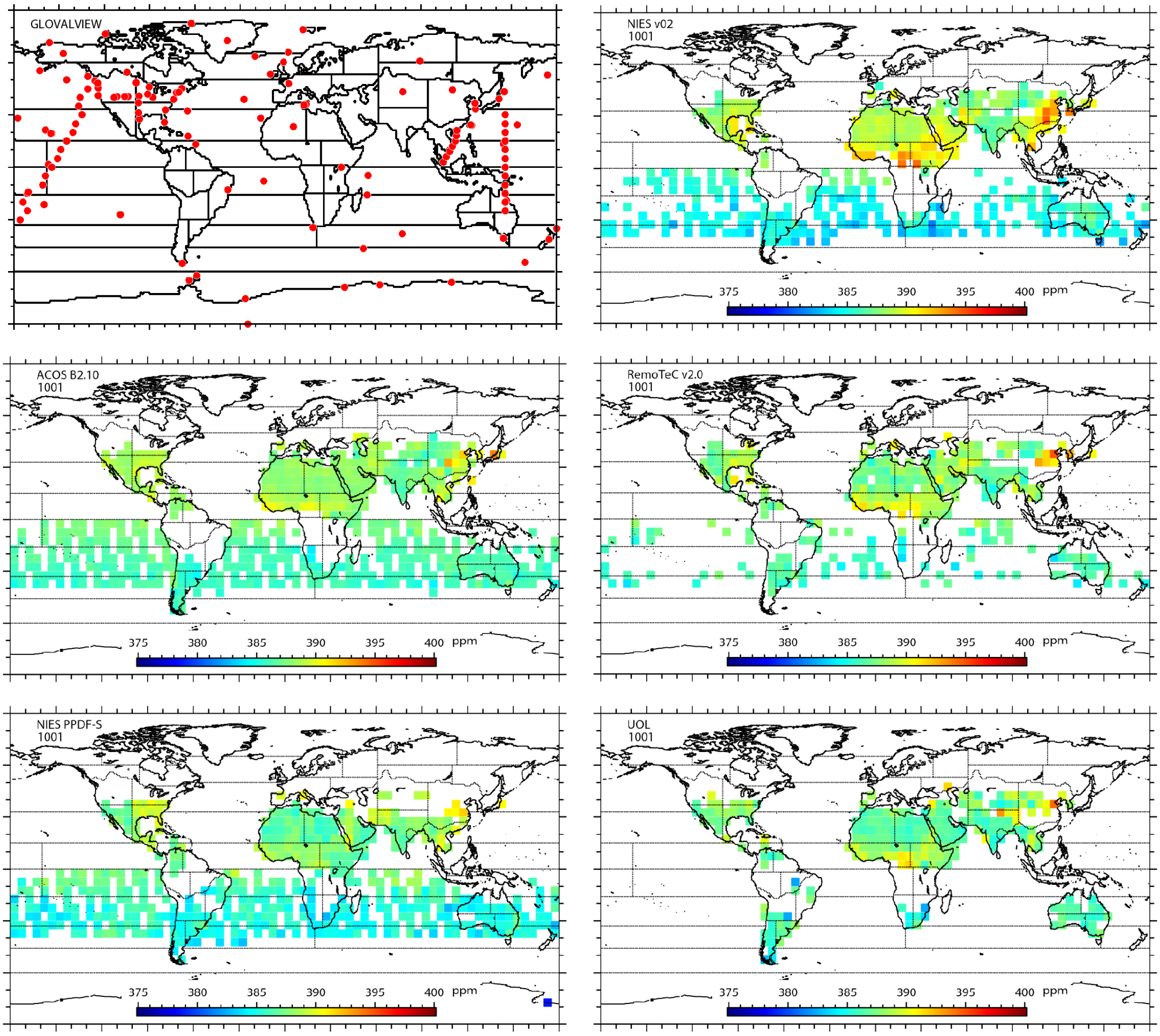

Figure 4S.22. Figure 4.7 for January 2010. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

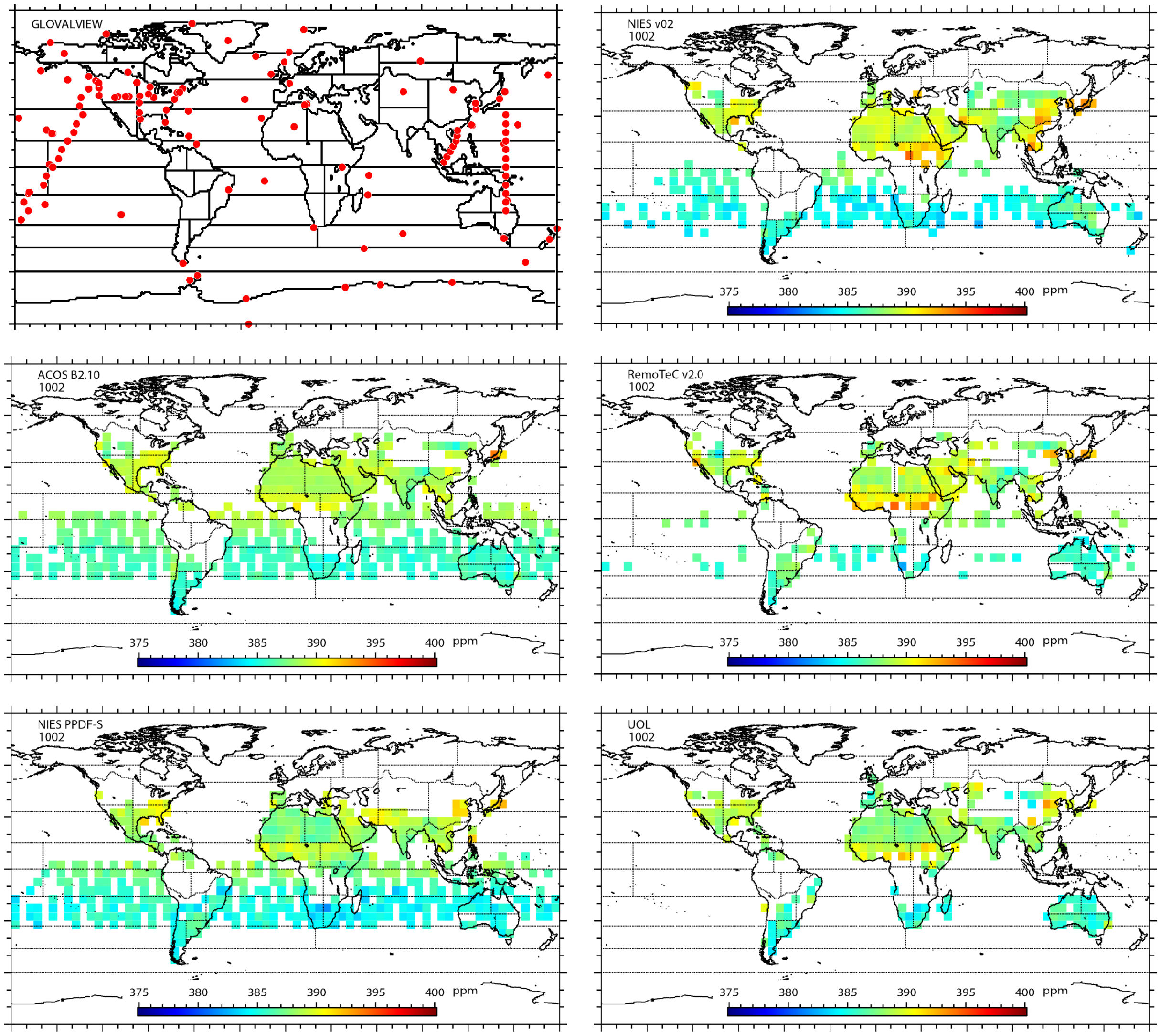

Figure 4S.23. Figure 4.7 for February 2010. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

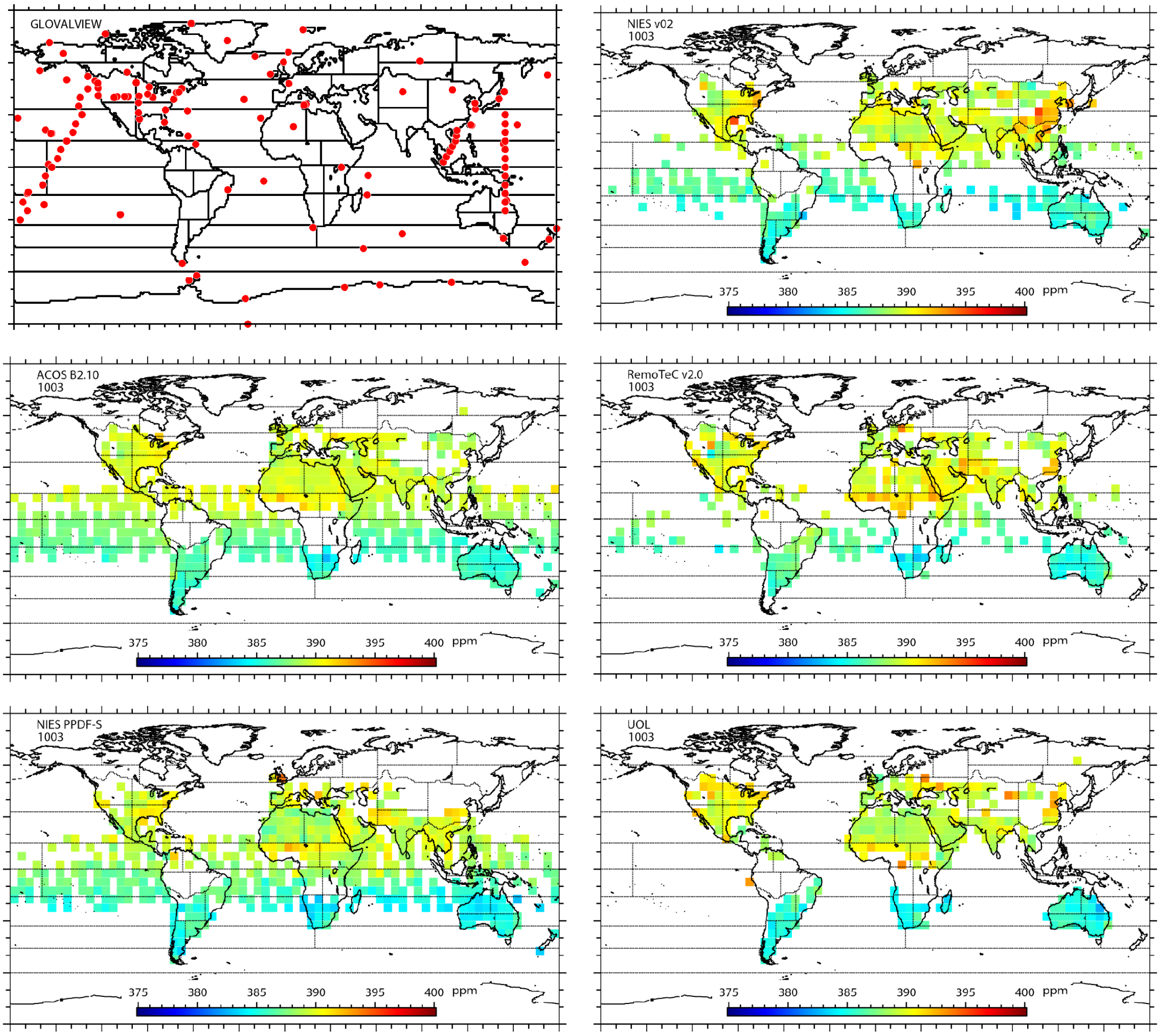

Figure 4S.24. Figure 4.7 for March 2010. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

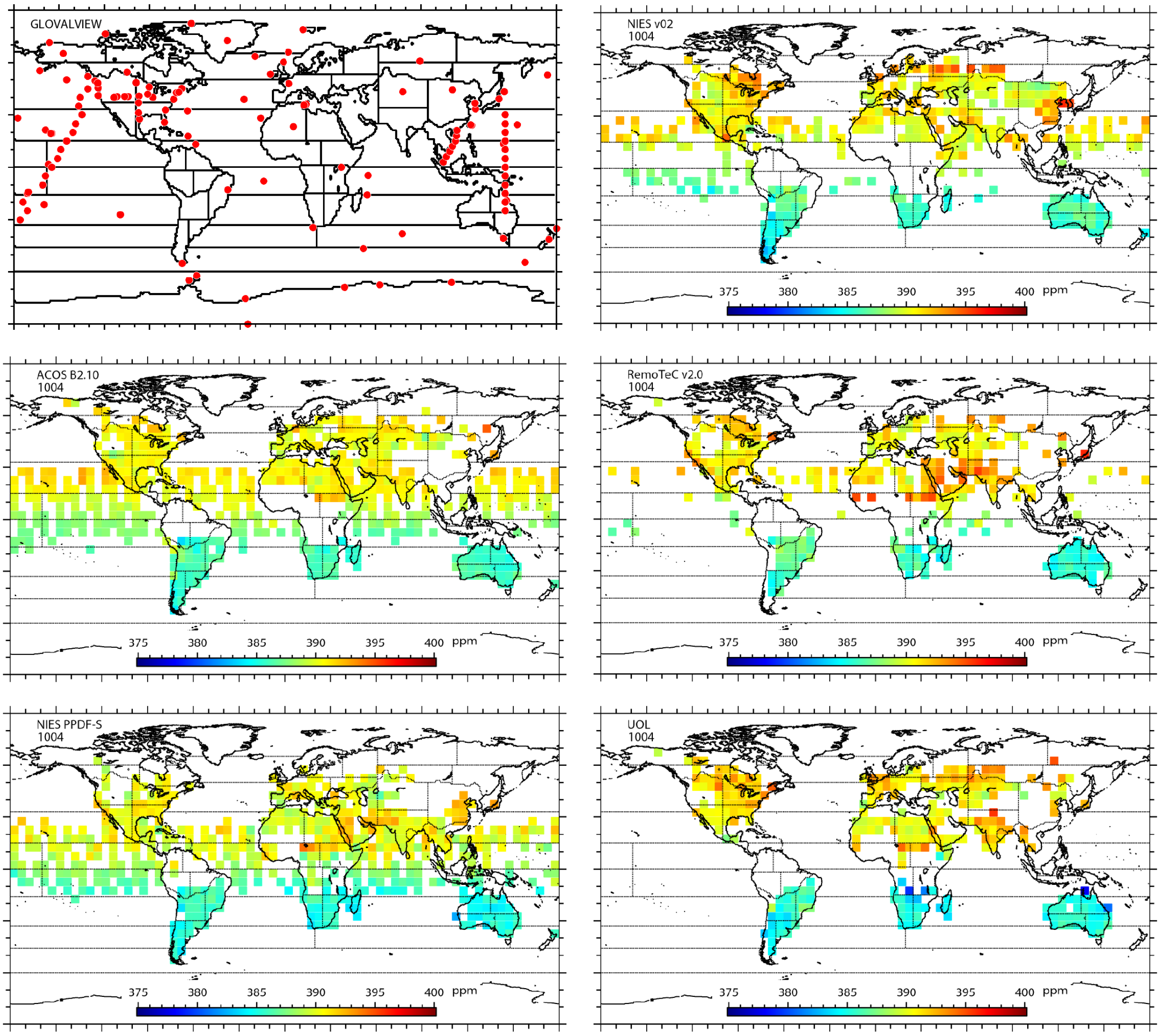

Figure 4S.25. Figure 4.7 for April 2010. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

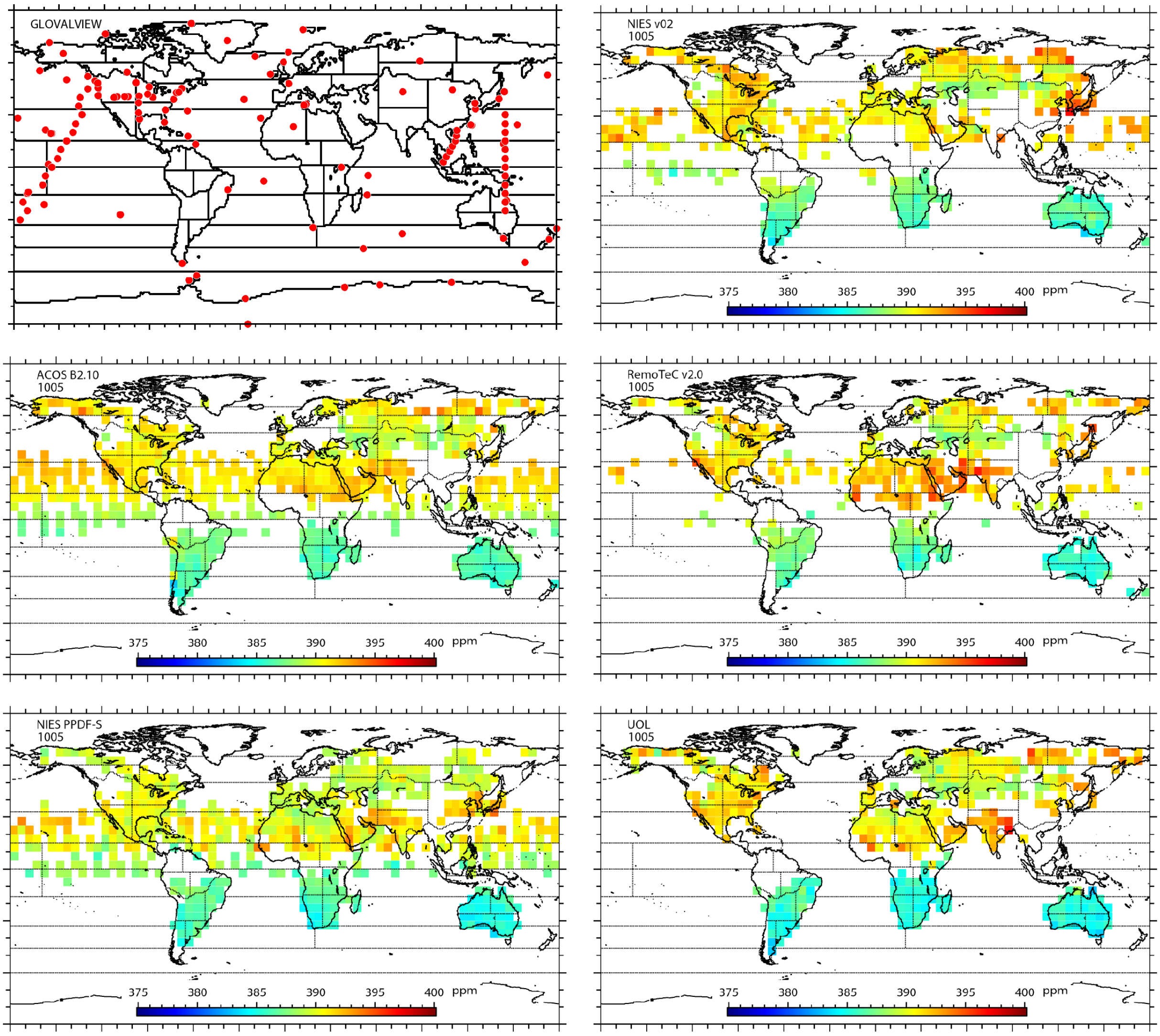

Figure 4S.26. Figure 4.7 for May 2010. Distributions of five 5°×5°- gridded X_{CO2} values retrieved with NIES v02 (upper right), ACOS B2.10 (middle left), RemoTeC v2.0 (middle right), PPDF-S (lower left), and UoL-FP v3G (lower right) algorithms.

CHAPTER 5

Impact of differences in spatial coverage of multiple GOSAT-based CO_2 datasets on regional flux estimates

This study was made possible through collaborating with the following researchers:

Robert J. Andres[1], Dmitry Belikov[2,3,4], Andrey Bril[2], Hartmut Boesch[5], Andre Butz[6], Otto Hasekamp[7], Sander Houweling[7,8], Makoto Inoue[2], Isamu Morino[2], Tomohiro Oda[9,10], Christopher W. O'Dell[9], Sergey Oshchepkov[2], Robert Parker[5], Makoto Saito[2], Osamu Uchino[2], Vinu Valsala[11], Tatsuya Yokota[2], Yukio Yoshida[2], and Shamil Maksyutov[2]

1 Oak Ridge National Laboratory, TN, USA

2 National Institute for Environmental Studies, Tsukuba, Japan

3 National Institute of Polar Research, Tokyo, Japan

4 Tomsk State University, Tomsk, Russian Federation

5 EOS Group, Department of Physics and Astronomy, University of Leicester, Leicester, UK

6 Karlsruhe Institute of Technology, Leopoldshafen, Germany

7 Netherland Institute for Space Research, Utrecht, The Netherlands

8 Institute of Marine and Atmospheric Research Utrecht, Utrecht, The Netherlands

9 Colorado State University, CO, USA

10 Global Monitoring Division, NOAA Earth System Research Laboratory, Boulder, CO, USA

11 Indian Institute for Tropical Meteorology, Pune, India

5.1. Introduction

The inference of regional CO_2 fluxes with the top-down approach, as introduced in Chapter 1, relies solely upon atmospheric CO_2 observations. As part of characterizing this inherent nature, several studies were conducted in the past to see the sensitivity of flux estimates to the choice of data-providing sites [e.g. Law et al., 2003; Yuen et al., 2005; Gurney et al., 2008] and to the expansion of surface monitoring networks over time [Bruhwiler et al., 2011]. These studies showed that changes in the geographical distribution of surface data have a large impact on regional-scale flux estimates.

With the advent of GOSAT in early 2009, CO_2 measurement by the surface monitoring networks is significantly augmented with the spaceborne X_{CO2} retrievals. As mentioned in Chapter 4, there exist five independent X_{CO2} retrieval datasets, and their precisions have been reported to be below 2 ppm level [Oshchepkov et al., 2013]. Where they coincide over land, the five X_{CO2} retrievals (bias corrected) were found to agree well within one standard deviation of about 1 ppm [Takagi et al., 2014]. Different from CO_2 measurements at fixed surface monitoring stations, success in the retrieval of satellite-based X_{CO2} is highly affected by the existence of light-scattering clouds and aerosols in the local sky, and therefore the chance that the X_{CO2} retrievals can be obtained again at the same location over the surface in the satellite's repeat cycle is not guaranteed. Also, in attempts to obtain better retrieval results, the five retrieval algorithms adopt different approaches in, e.g., modeling the vertical distribution of clouds and aerosols and screening low-quality X_{CO2} retrievals. Thus, it is highly possible that the spatial distributions of X_{CO2} retrievals yielded by each of the five algorithms differ from one another.

Takagi et al. [2014], in their study on the influence of differences in five independent GOSAT X_{CO2} datasets on flux estimates (content presented in Chapter 4), concluded that large spread among five fluxes estimated for temperate Asia regions could be linked to differences in spatial data coverage by each of the five X_{CO2} datasets. They also suggested that the flux spread could be more pronounced if individual, "single-shot" X_{CO2} values (as opposed to gridded and monthly-averaged values) were used in the estimation.

Here, I investigate further this previously addressed topic by shedding light on the extent that the differences in the X_{CO2} data spatial coverage alter constraints on flux estimates. For this, I estimated monthly fluxes for the same 64 source regions as explained in Chapter 4, but this time I used single-shot X_{CO2} values as stored in each of the five datasets; also, the X_{CO2} values were not used in combination with surface-based CO_2 data as in the previous experiment to isolate their contributions to the flux estimation. My focus here was directed onto temperate Asia, in particular its north eastern region that cover Japan, eastern China and the Korean Peninsula (Region 32) where the spread among the five flux estimates in the previous experiment was found to be large. For comparison, constraints on five regional fluxes estimated for this region were quantified and visualized by using two diagnoses, response function and resolution kernel (explained in Sections 5.2.2 and 5.2.3).

5.2. Data and method

5.2.1 X_{CO2} retrieval datasets

The five X_{CO2} retrieval datasets considered here are as follows (the updated versions of the datasets listed in Chapter 4 were used here): NIES v02.11, PPDF-S v02.11, ACOS B3.4, RemoTeC v2.11, and UoL-FP v4. The biases in the X_{CO2} values stored in these datasets were corrected by the individual research groups, using linear regressions that correlate variabilities in X_{CO2} values and selected retrieval parameters [Wunch et al., 2011a; Guerlet et al., 2013; Cogan et al., 2012; Inoue et al., in preparation]. The data over land and ocean are stored in these datasets, except that UoL-FP v4 comes with land values only. To perform the flux inter-comparison under an equal condition, I used only the land retrievals stored in each of the datasets. Also, since the end of the time period that each of the datasets covers is not the same, I used the retrieval values over a period from June 2009 to March 2011 for inverse modeling (25 months). Among those analyzed months, the focus here was directed onto year 2010.

For estimating random errors associated with the X_{CO2} datasets considered, I compared the five bias-corrected X_{CO2} datasets over year 2010 against reference data obtained at the TCCON observational sites. Each GOSAT X_{CO2} retrieval found within a $\pm 2°$ grid box centered at each of 11 selected TCCON sites was compared with TCCON X_{CO2} data that were averaged over ± 30 min. of local GOSAT overpass time. It turned out that the standard deviations (SD) of GOSAT-TCCON differences, averaged over the one-year analyzed period, ranged from 1.6 ppm (PPDF-S v02.11) to 2.0 ppm (RemoTeC v2.11 and UoL-FP v4) (see Table 5.1). These are about twice the global, one-year mean of the SDs of five collocated X_{CO2} values over land (0.8 ppm; sample distributions of collocated X_{CO2} SDs are found in Figure 5.1). This suggests that the agreement among the five X_{CO2} retrievals is well met within the range of their random errors and that the focus of the satellite-based inversion inter-comparison can now be directed onto the differences in the data spatial coverage.

To contrast the satellite-based inversion results with one based on data from existing surface monitoring networks, I obtained another estimate using GV (2012 issue). Data from 212 monitoring locations were selected (Figure 5.2), and they were monthly-averaged when used in the inversion.

5.2.2 Inverse modeling setup

The inverse modeling system and a priori flux datasets used here are the same as the ones used in Chapter 4, except that the individual X_{CO2} values were used (thereby not gridded nor monthly-averaged) in the inversion. The model-observation mismatch errors for X_{CO2} retrievals, stored in the diagonal elements of square matrix $\mathbf{C_D}$, were set at the sum of 2 ppm random error (Section 5.2.1) and the forward modeling uncertainty of 1 ppm as reported by Belikov et al. [2013] (3 ppm total). The values for GV data were taken from residual SDs that are recorded in the GV 2012 dataset. The minimum mismatch error for GV data was set at 0.3 ppm. For this experiment, I ran the system to estimate monthly fluxes for the 64 source regions over the 25 modeling months.

5.2.3 Response functions

As introduced earlier in Chapter 2, for each of the monthly regional fluxes estimated in this analysis, a concentration simulation was performed in which a unit emission of 1 GtC region^{-1} yr^{-1} was released from that region for one month and transported forward until the end of the simulation period to sample responses at the time and location of every X_{CO2} retrieval. The spatial pattern of the 1 GtC region^{-1} yr^{-1} unit emission for each of the 42 land source regions was defined as that of 31-yr-mean (1980-2010) net primary productivity estimated by VISIT terrestrial biosphere process model. No spatial patterns were given to the unit emissions for the 22 ocean basins (spatially uniform). The sampled responses, or the response functions, were recorded in the columns of matrix **G**, which functions as a linear operator that relates concentrations with regional flux magnitudes. The responses in matrix **G** represent the degree of the contribution of individual X_{CO2} retrievals to constraining regional monthly fluxes.

The magnitude of a response to a unit emission from a region, as stored in matrix **G**, is dependent on 1) the horizontal pattern of the unit emission, 2) atmospheric transport (which changes with time and space), and 3) the time and location of X_{CO2} retrieval. The unit emission patterns vary from region to region, and in some regions there exist highs and lows in their emission patterns owing to the distribution of land cover types. Such a contrast is clearly seen in the unit emission pattern for Region 32, and is shown in Figure 5.3. The contrast seen over the continental Region 32 comes from its land cover type that changes from its northern part (grasslands and barren fields) to the southern part (mostly mixed forests). Because of this north-to-south contrast, responses sampled closer to the emission sources can be higher than those away from the sources. This suggests that the response magnitudes, which are related to constraints on regional fluxes, are dependent on where the X_{CO2} retrievals cover and how many of them exist in and around a region of interest.

5.2.4 Resolution kernel

A convenient diagnostic to show the degree to which observations constrain the estimated fluxes is the resolution kernel [Tarantola, 1987; Menke, 1989; Bruhwiler et al., 2011]. It is a square matrix whose rank is equal to the number of individual fluxes estimated, and is derived from the error covariance matrix associated with the a posteriori flux estimates,

$$\mathbf{C'_M} = \mathbf{C_M} - \mathbf{C}_\mathrm{M}\,\mathbf{G}^\mathrm{t}(\mathbf{G}\,\mathbf{C}_\mathrm{M}\,\mathbf{G}^\mathrm{t} + \mathbf{C}_\mathrm{D})^{-1}\mathbf{G}\,\mathbf{C_M}, \qquad (2\text{-}13)$$

or

$$\mathbf{C'_M} = (\mathbf{I} - \mathbf{R})\,\mathbf{C_M}, \qquad (5\text{-}1)$$

where **R** is given as

$$\mathbf{R} = \mathbf{C}_\mathrm{M}\,\mathbf{G}^\mathrm{t}(\mathbf{G}\,\mathbf{C}_\mathrm{M}\,\mathbf{G}^\mathrm{t} + \mathbf{C}_\mathrm{D})^{-1}\mathbf{G}\,, \qquad (5\text{-}2)$$

the resolution kernel (RK). RK is equivalent to the averaging kernel in the retrieval of X_{CO2} values. Equation 5-1 suggests that as **R** comes close to **I** (identity matrix; diagonal elements are unity), $\mathbf{C'_M}$ approaches 0; such a posteriori flux estimates can be considered as well resolved by the observations. Also, Equation 5-2 indicates the dependence of **R** on the linear operator matrix **G** whose row size reflects the availability of observational data for resolving the regional fluxes. The row size and the magnitude of the elements in the columns of **G** together represent how well the retrieval datasets (or the surface-based data) can resolve regional fluxes. I will use this diagnostic to see quantitatively how the differences in the spatial coverage by the five retrieval datasets bring changes to the constraints on the regional fluxes.

5.3. Results

Presented in Figure 5.4 is the time series of fluxes estimated for Region 32 for 2010. The seven solid lines in the figure show the following estimates: a priori (its uncertainty is shown with shade), GV-only, NIES, ACOS, PPDF-S, RemoTeC, and UoL-FP. The values shown are in gC m^{-2} day^{-1}, and are without anthropogenic emissions. The five GOSAT-based flux estimates agreed well after August; large disagreements are found from February to August. The annual regional total flux in GtC yr^{-1} thus turned out to be variable: 0.9 (NIES), -0.8 (ACOS), -0.7 (PPDF-S), 0.8 (RemoTeC), and -0.7 (UoL-FP). The smallest and largest spreads (maximum value minus minimum value among the five flux estimates in a month) are found in September (0.5 gC m^{-2} day^{-1}) and April (1.9 gC m^{-2} day^{-1}), respectively. Below, I will present the results for these two contrasting months.

The circles in the upper panels of Figure 5.5 show the horizontal distribution of the locations of X_{CO2} retrievals that contributed to the estimation of September 2010 fluxes (characterized with the small flux spread). The color in each circle denotes the magnitude of the response (Section 5.2.3) sampled at the time of GOSAT measurement. Presented in Figure 5.6 are the distributions of monthly-mean responses on a 2.5-degree grid for April and September 2010. The figure shows the prevailing trend of atmospheric tracer transport (the responses) within and around Region 32 on a monthly timescale. Over the continental Region 32, there is an ellipse-shaped region of high responses whose center is located over the locations of high surface emissions. The circles shown in Figure 5.5 that are close to the high response center are colored in warmer colors in both months

(April and September). The extent of the high response area varies owing to seasonal changes in atmospheric transport, therefore the distribution of the colored circles found in Figure 5.5 also changes with season. It can be seen in Figure 5.6 that below 0.4 ppm the light-green outer edge of the ellipse blend quickly into the background (blue color). Here in this analysis, I use this 0.4 ppm boundary as a threshold for distinguishing significant or "influential" responses from those at the background level and characterizing each of the five X_{CO2} spatial distributions.

The spatial coverage by each X_{CO2} dataset for September 2010 is not exactly identical to one another, yet each dataset covers well the higher response grids from near the high center to the outer perimeter. I counted the number of individual measurement locations shown on Figure 5.5 at which sampled responses are greater than the significance threshold (0.4 ppm). Then I calculated the averages of the following values at those locations considered as influential in constraining the fluxes: 1) responses, 2) retrieved X_{CO2} concentrations, 3) a posteriori (optimized) X_{CO2} concentrations, and 4) differences between the retrieved and the optimized (residuals). These values for each of the five retrieval dataset are listed in Table 5.2A (September case). The number of influential measurement locations varied from 53 (ACOS and PPDF-S) to 118 (RemoTeC). The averages of retrieved X_{CO2} concentrations for all the cases were found to be around 387 ppm (386.9 $\pm$ 0.3 ppm), and the a posteriori X_{CO2} concentrations were very close to that range (387.1 $\pm$ 0.3 ppm). The narrow retrieved X_{CO2} range supports the small spread among the five flux estimates for this month.

The RK values for the September flux estimation are shown in the upper panel of Figure 5.7. The diagonal RK value for the GV-only case, indicated with region ID 32 on the horizontal axis, was 0.73 in both the September and April cases. This value was similar to one obtained by Bruhwiler et al. [2011] for broader Temperate Asia region (0.7; area equal to Regions 29-32 combined) using a surface network configuration for the year 2000. The diagonal RK values for the five GOSAT-based cases are all above 0.9 (range: 0.94 (PPDF-S) – 0.98 (UoL-FP)), signifying that the fluxes were resolved by the X_{CO2} retrievals better than the GV data. The difference between the GV-only and GOSAT-based RKs can indicate the amount of extra information that can be supplied by the wide-covering X_{CO2} retrievals that are larger in number density but less precise than the surface-based data (minimum uncertainty of 3 ppm specified for X_{CO2} retrievals as opposed to ~0.3 ppm for GV data). The differences seen in the GOSAT-based RK values are found to be reflective of the differences in the number of influential X_{CO2} retrievals counted. The off-diagonal RKs found elsewhere (RK values in Figure 5.7 other than one for Region 32 (the diagonal RK)) are all below 0.3. RK values at ~0.3 level were found in the GV-only case for remote regions such as tropical America (Regions 9-12) whose fluxes were inferred from data collected in distant regions (no GV sites within these regions; see Figure 5.2). A sample RK for Region 09 (tropical America SW) is shown in Figure 5S.1 (notice the GV-only RK values in red that are below 0.3 throughout the year). The low off-diagonal RKs found in the September case (Figure 5.7) suggest that the Region 32 fluxes are well distinguished from the estimates for the neighboring regions.

The lower panels of Figure 5.5 show the response distributions for the April 2010 estimation (characterized with large flux spread). Corresponding data numbers and average values are listed in Table 5.2B (April case). The horizontal extent of the higher response grids (>~0.4 ppm) is much more limited than that of the September case; as shown in Figure 5.6, the location of the high center is about the same, but the northern edge of the higher response field does not reach 40° N. The spatial distributions of the measurement locations over the region differ largely from dataset to dataset; those of UoL-FP and RemoTeC are quite contrasting, particularly in the south. The majority of the measurement locations of the five X_{CO2} datasets were found in the northern part of continental Region 32 (away from the strong sources), and therefore their responses are low. Only a few X_{CO2} retrievals located within or near the higher response field were counted to be influential; the number ranged from 5 (ACOS) to 22 (NIES) (see Table 5.2B). These numbers are much smaller than those found in the September case (Table 5.2A), suggesting increased cloudiness and/or atmospheric aerosol loading in the southern part of the continental Region 32 in this month. The differences in the total number of measurement also suggest that each of the five retrieval algorithms screens the satellite measurements and retrieval results quite differently.

The averages of the a posteriori X_{CO2} concentrations differed one to another in the April case (range: 393.6-395.0 ppm). The mean a posteriori concentrations for the NIES and RemoTeC cases are about 395 ppm, and those for the remaining cases are all below that level (<394 ppm). The monthly flux estimates associated with the higher mean concentration cases (NIES and RemoTeC) are >1.5 gC m^{-2} day^{-1}, and the others turned out to be below 0.7 gC m^{-2} day^{-1}. This concentration-flux relation is shown in Figure 5.8 (right). Note here that this trend is not seen in the September case (Figure 5.8 left).

The diagonal RK values for the April case (Figure 5.7, bottom panel) are reflective of the numbers of the influential measurements that are much smaller than those seen in the September case (>50). The individual RK values are all smaller than the corresponding September values, and varied from 0.79 (ACOS) to 0.93 (UoL-FP). The diagonal RK for ACOS in this case is nearly comparable to that for GV-only (0.73).

For the Region 32 monthly cases in 2010, I found a clear correlation (r = 0.6) between the level of agreement in the fluxes (SD of five fluxes) and the variability in the diagonal RK values (SD of five RKs), and it is shown in Figure 5.9. Other regional cases in which clear correlations were found between flux SD and RK SD (e.g. Regions 16, 22, 28, and 30) showed variability in the data spatial coverage by the influential X_{CO2} retrievals similar to that found in Region 32.

5.4. Concluding remarks

Based on the fact that the recent versions of five bias-corrected GOSAT X_{CO2} retrievals over land agree reasonably well within the range of their random errors (Section 5.2.1 and Figure 5.1), I investigated how the differences in spatial coverage by the five retrieval datasets can alter constraints on regional monthly flux estimates. I found, based on the results obtained for the Temperate Asia NE region (Region 32), that constraint on a regional flux is dependent on how the influential X_{CO2} retrievals are spatially distributed. I showed quantitatively the alteration in flux constraint, using the resolution kernel (RK), a diagnostic for indicating the degree to which a set of observations constrain flux estimates (defined in Section 5.2.4). April 2010, one of the two focused months, was the month in which the spread between the largest and smallest flux estimates was 1.9 gC m^{-2} day^{-1} (Figure 5.4), and I observed in this case that the data spatial coverage differ largely from one dataset to another (Figure 5.5, bottom panels). I saw that this spread was signified in the larger variability in RK values (Figure 5.7, bottom panel) and in the averages of retrieved X_{CO2} concentrations that were classified as influential to the regional flux estimation (Table 5.2B). I found a clear correlation between the level of agreement in fluxes (SD of five flux estimates) and the variability in RK (SD of five RK values) in this region in this analyzed year, and also in other regions where the coverage patterns differ largely from one dataset to another.

The April 2010 case may also represent other regional flux estimation cases in which data coverage patterns change with season, or perhaps from year to year for various reasons (e.g., changes in local clear-sky probabilities, infrequent large-scale forest fire events, etc.). In those cases, flux estimates with diagonal RK values that vary largely with time may need to be analyzed carefully as they potentially contain uncertainties associated with changing data number density and coverage.

References

Belikov, D. A., et al. (2013), Simulations of column-averaged CO_2 and CH_4 using the NIES TM with a hybrid sigma-isentropic (σ-θ) vertical coordinate, *Atmos. Chem. Phys.*, **13**, 1713-1732, doi:10.5194/acp-13-1713-2013.

Bruhwiler, L. M. P., A. M. Michalak, and P. P Tans, (2011), Spatial and temporal resolution of carbon flux estimates for 1983–2002, *Biogeosciences*, **8**, 1309-1331, doi:10.5194/bg-8-1309-2011.

Cogan, A. J., et al. (2012), Atmospheric carbon dioxide retrieved from the Greenhouse gases Observing SATellite (GOSAT): Comparison with ground-based TCCON observations and GEOS-Chem model calculations, *J. Geophys. Res.*, **117**, D21301, doi:10.1029/2012JD018087.

GLOBALVIEW-CO2: Cooperative Atmospheric Data Integration Project - Carbon Dioxide. NOAA ESRL, Boulder, Colorado, 2012.

Gurney, K. R., D. Baker, P. Rayner, and S. Denning (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.

Inoue M. et al., Bias correction and evaluation of GOSAT SWIR X_{CO2} and X_{CH4} using TCCON data and aircraft measurements, in preparation

Law, R. M., Y.-H. Chen, K. R. Gurney and TransCom 3 Modellers (2003), TransCom 3

CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, *Tellus*, **55B**, 580–595. doi: 10.1034/j.1600-0889.2003.00053.x.

Menke, W. (1989), Geophysics Data Analysis: Discrete Inverse Theory, vol. 45, International Geophysics Series, Academic Press, San Diego, revised edition.

Oshchepkov, S., et al. (2013), Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space. Part 2: Algorithm intercomparison in the GOSAT data processing for CO_2 retrievals over TCCON sites, *J. Geophys. Res. Atmos.*, **118**, 1493–1512, doi:10.1002/jgrd.50146.

Takagi, H. et al. (2014), Influence of differences in current GOSAT X_{CO2} retrievals on surface flux estimation, *Geophys. Res. Lett.*, **41**, 2598–2605, doi:10.1002/2013GL059174

Tarantola, A. (1987), Inverse Problem Theory, Elsevier, New York.

Wunch, D., et al. (2011), A method for evaluating bias in global measurements of CO_2 total columns from space, *Atmos. Chem. Phys.*, **11**, 12317-12337, doi:10.5194/acp-11-12317-2011.

Yuen, C.-W., K. Higuchi, and TransCom-3 Modellers (2005), Impact of Fraserdale CO_2 observations on annual flux inversion of the North American boreal region, *Tellus,* **57B**, 203-209.

Tables

Table 5.1. SD of GOSAT-TCCON differences averaged over 2010

	NIES v02.11	**ACOS B3.4**	**PPDF-S v02.11**	**RemoTeC v2.11**	**UoL-FP v4**
Number of land retrievals (2010, 1 yr.)	59316	59424	79189	53314	86815
Global mean and SD of GOSAT-TCCON differences (ppm)	0.2±1.9	0.3±1.6	0.4±1.7	0.4±2.0	0.4±2.0

Table 5.2. The number of influential measurements counted in Region 32 (red) and averages of concentrations. The four concentrations shown (response, retrieved X_{CO2}, a posteriori X_{CO2}, and retrieved-a posteriori residual) are the averages of values considered as influential (> 0.4 ppm threshold).

A: September 2010

	NIES v02.11	ACOS B3.4	PPDF-S v02.11	RemoTeC v2.11	UoL-FP v4
Num. of data counted*	79 / 383	53 / 528	53 / 362	118 / 286	108 / 603
Avg. Response (ppm)	0.6	0.6	0.7	0.7	0.7
Avg. Retrieved (ppm)	387.4	387.1	386.8	386.9	386.6
Avg. Posterior (ppm)	387.6	387.3	387.1	387.2	386.6
Avg. Residual (ppm)	0.2	0.2	0.3	0.3	0.0

B: April 2010

	NIES v02.11	ACOS B3.4	PPDF-S v02.11	RemoTeC v2.11	UoL-FP v4
Num. of data counted*	22 /222	5 / 211	9 / 84	9 / 83	18 / 267
Avg. Response (ppm)	0.6	0.5	0.6	0.6	0.8
Avg. Retrieved (ppm)	395.1	394.6	392.3	394.9	393.6
Avg. Posterior (ppm)	395.0	393.8	393.6	394.7	393.9
Avg. Residual (ppm)	-0.1	-0.8	1.3	-0.2	0.3

* The number in black is the total number of measurement found in Region 32 in the month. The number in red is the number of influential measurement (> 0.4 ppm threshold; see Section 3).

Figures

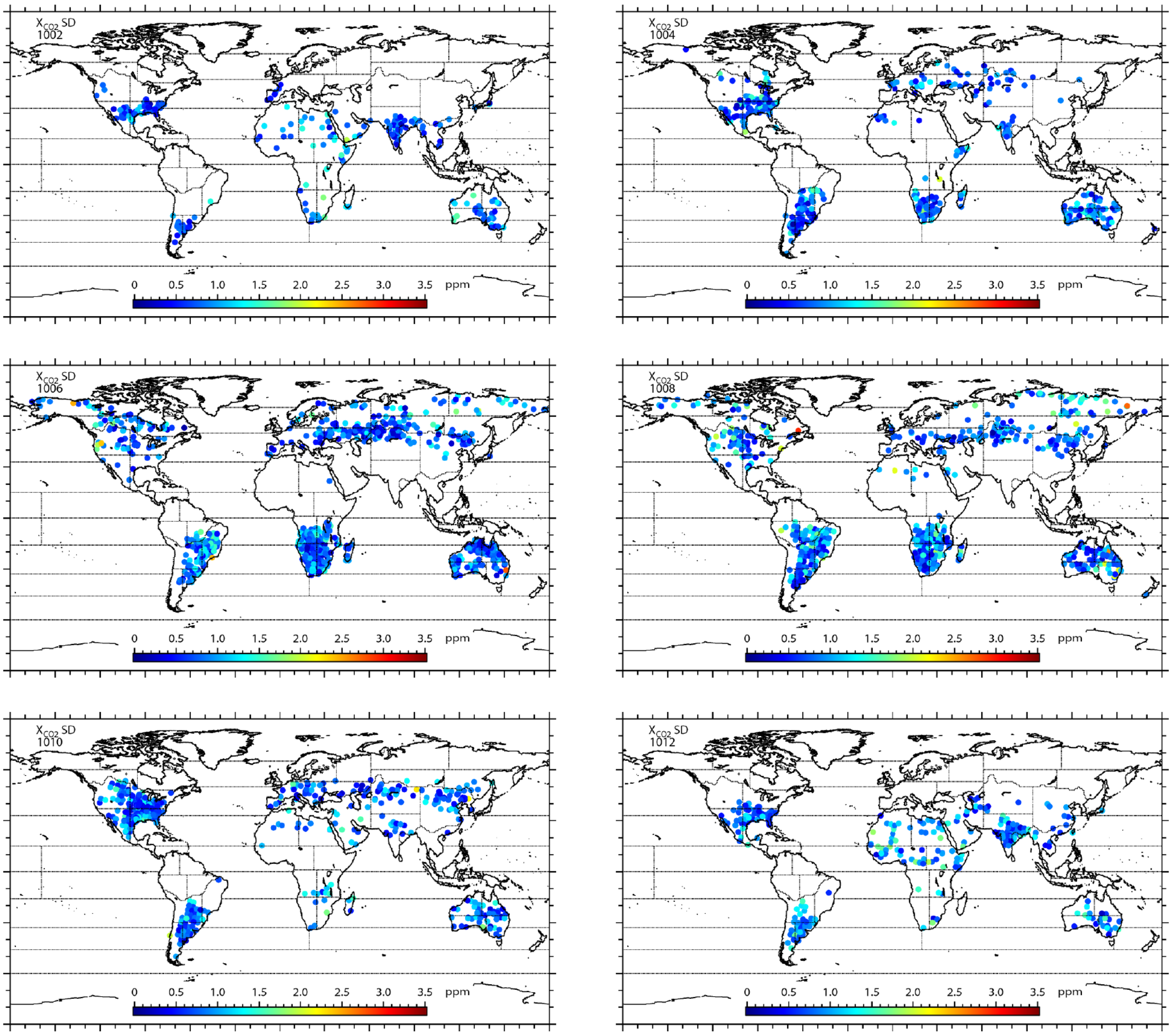

Figure 5.1. Standard deviation of five collocated X_{CO2} retrievals. Values for the even-numbered months in 2010 are shown (indicated in YYMM format).

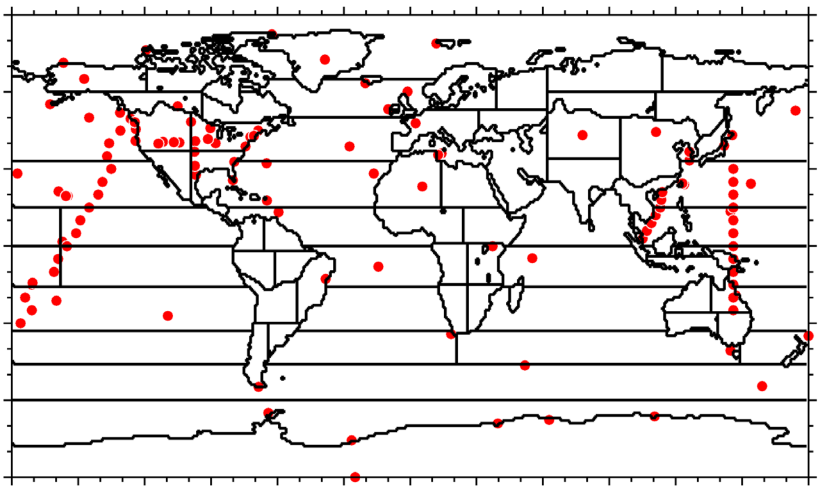

Figure 5.2. Locations of GV monitoring stations selected in this study (212).

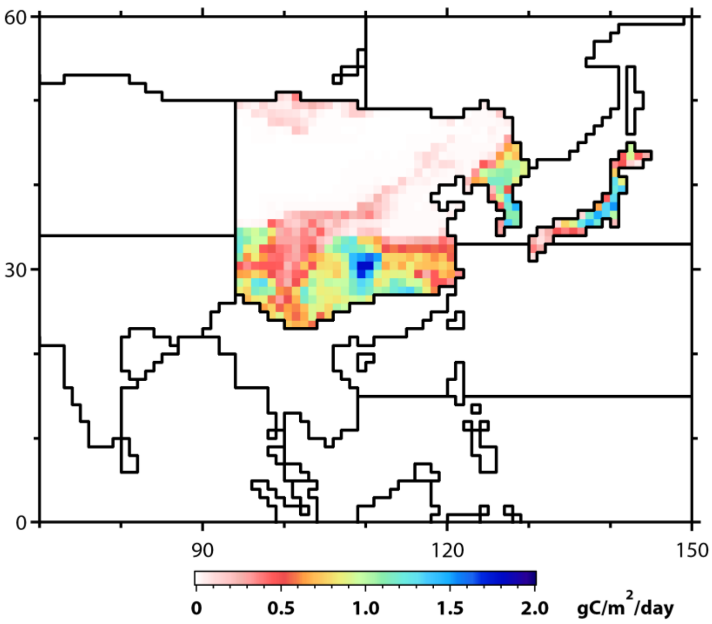

Figure 5.3. Unit emission pattern for Region 32 (Temperate Asia NE region that covers eastern China, part of Mongolia, the Korean Peninsula, and Japan).

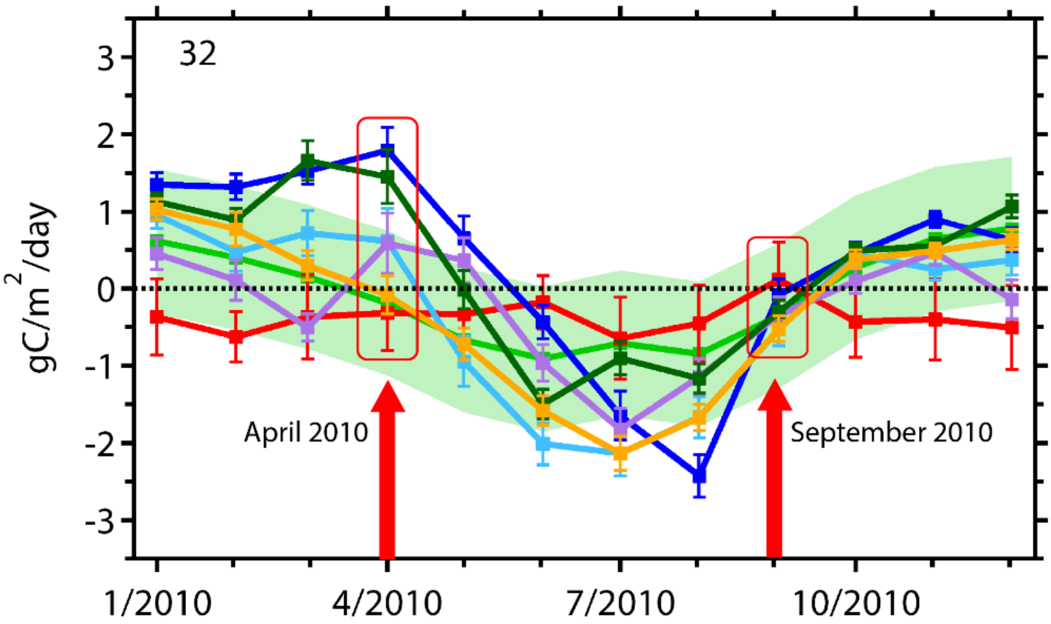

Figure 5.4. Time series of monthly fluxes estimated for Region 32 for 2010. The seven solid lines in the figure show the following: a priori (light green; its uncertainty is shown with green shade), GV-only (red), NIES (blue), ACOS (light blue), PPDF-S (purple), RemoTeC (green), and UoL-FP (orange). The error bar indicates a posteriori uncertainty.

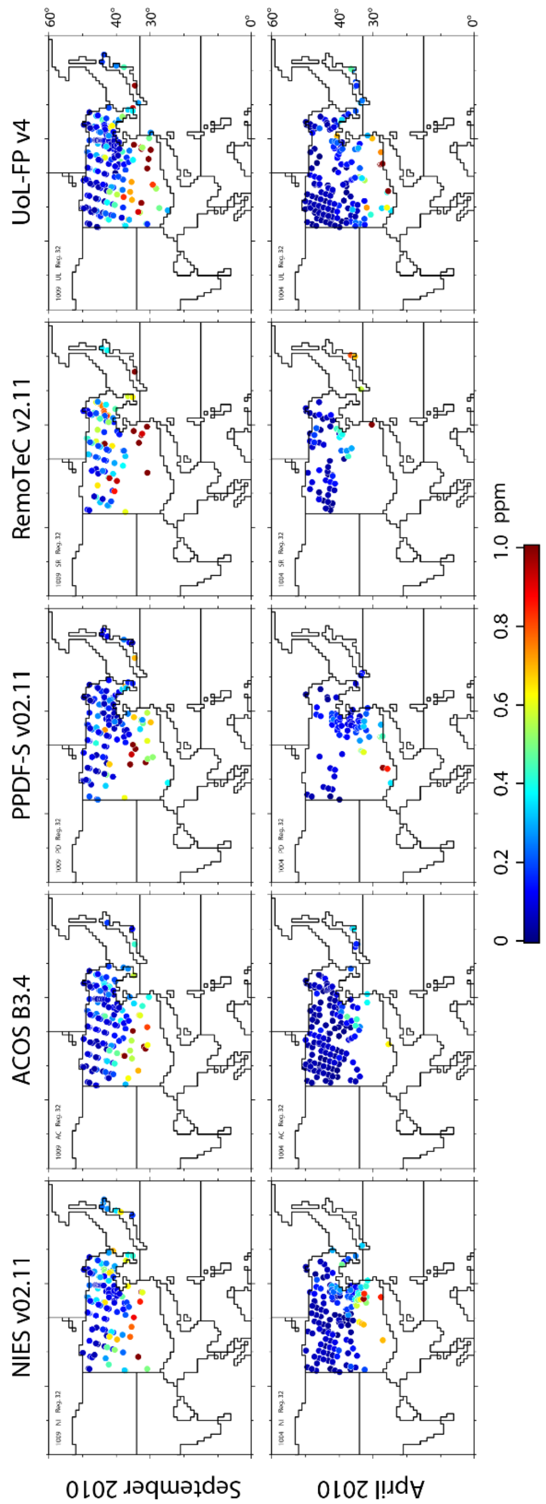

Figure 5.5. Horizontal distribution of the locations of X_{CO2} retrievals that were used for the flux estimation. The color in each circle denotes the response sampled at the time of GOSAT spectral measurement. Upper row: September 2010. Lower row: April 2010.

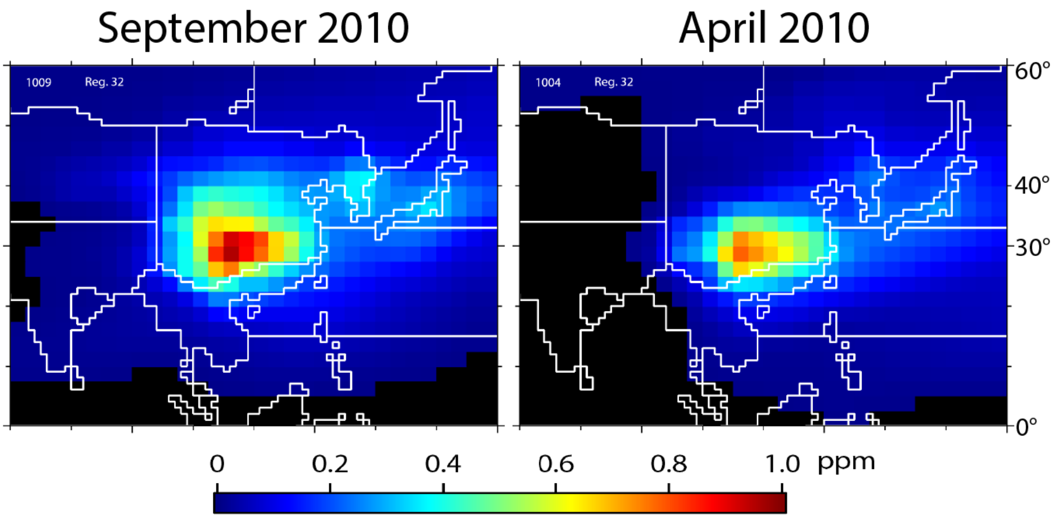

Figure 5.6. Distribution of monthly-mean responses on a 2.5°×2.5° grid in Region 32 for April and September 2010.

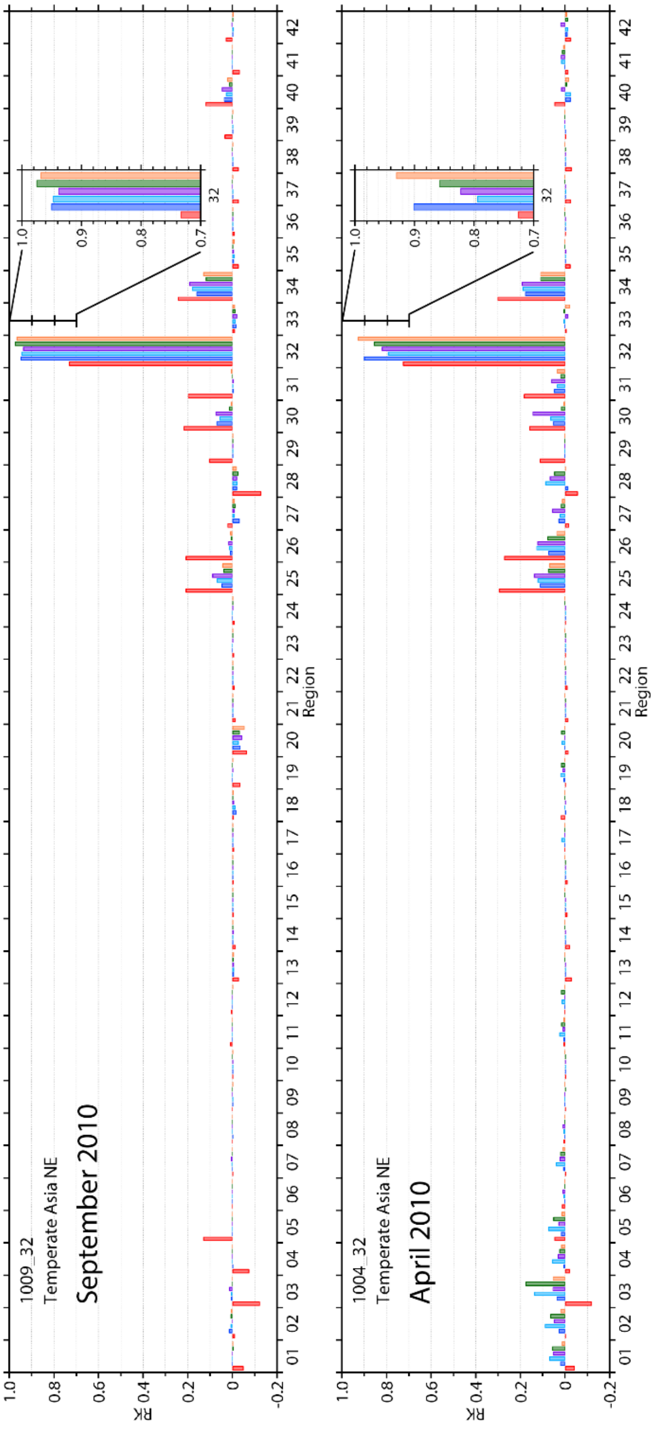

Figure 5.7. Resolution Kernel for Region 32 for the September (upper panel) and April (lower panel) flux estimation. Red: GV-only. Blue: NIES. Light blue: ACOS. Purple: PPDF-S. Green: RemoTeC. Orange: UoL-FP.

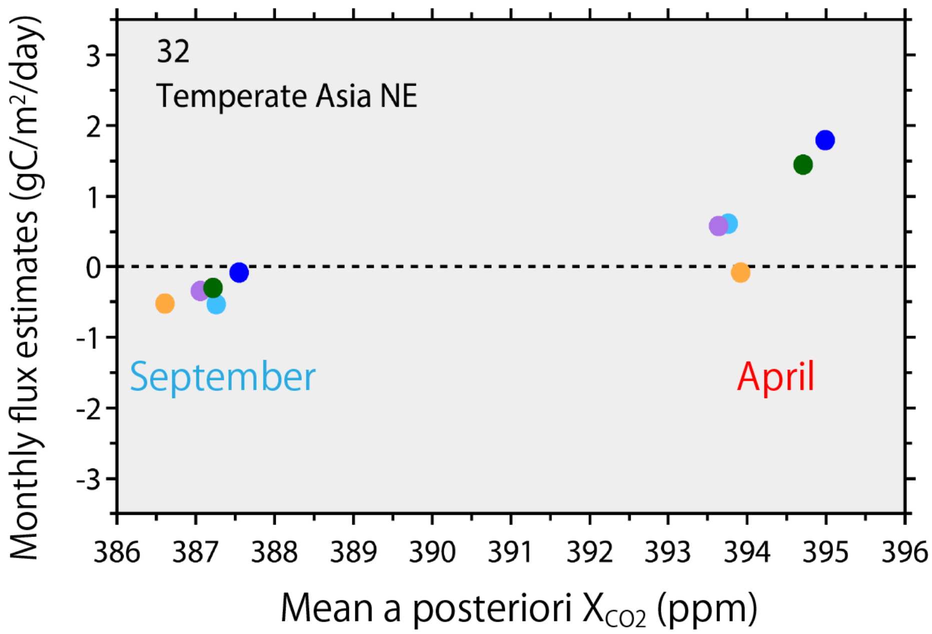

Figure 5.8. Monthly flux estimates vs. corresponding mean a posteriori X_{CO2} concentrations. Values for the September (left) and April (right) cases are shown. Blue: NIES. Light blue: ACOS. Purple: PPDF-S. Green: RemoTeC. Orange: UoL-FP.

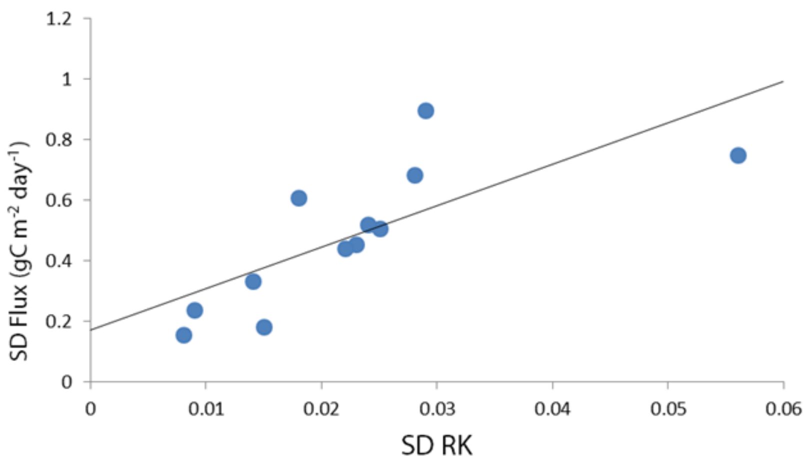

Figure 5.9. Standard deviation (SD) of five flux estimates vs. SD of five diagonal RK values. Values for Region 32 in the analyzed 12 months in 2010 are shown.

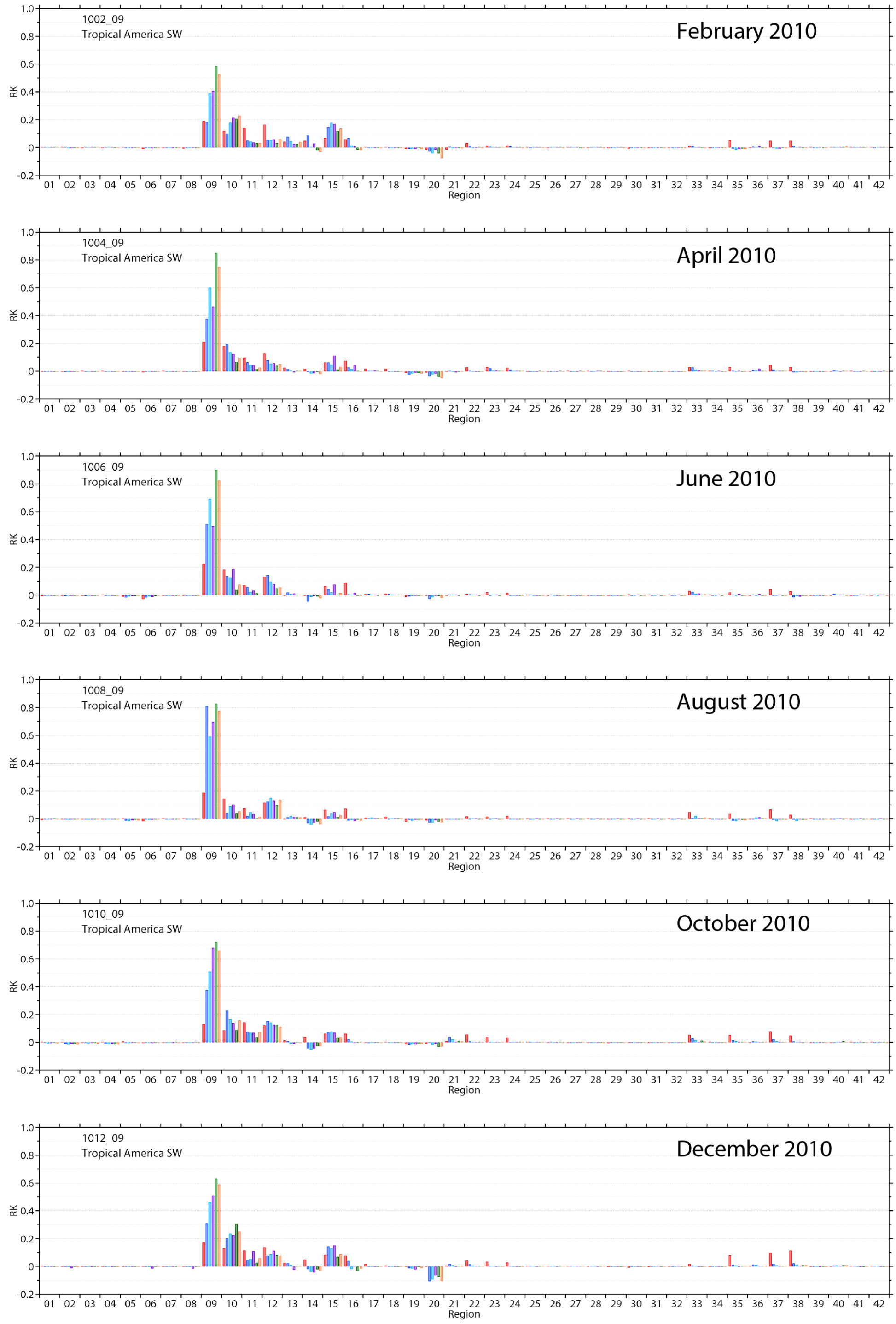

Figure 5S.1. RK for tropical America region (Region 09). Values for even-numbered months are shown. Red: GV-only. Blue: NIES. Light blue: ACOS. Purple: PPDF-S. Green: RemoTeC. Orange: UoL-FP. Notice that the GV-only RKs in red are all below 0.3 level throughout the year.

CHAPTER 6

Summary and perspective on future studies

The advent of GOSAT in 2009 has brought a new era in the estimation of surface CO_2 fluxes, providing researchers with an unprecedented amount of global CO_2 concentration data than ever made available for global carbon cycle studies. A top-down, inverse modeling system and its subsystems developed were used to estimate monthly fluxes for 64 global regions from both GLOBALVIEW surface-based CO_2 data and GOSAT X_{CO2} retrievals for the first time in the world. It was found that the addition of grid-aggregated monthly GOSAT X_{CO2} retrievals to the existing surface-based data brings reductions in the a posteriori flux uncertainties as much as about 60 % during the analyzed one year (June 2009 - May 2010) (first in the world in evaluating the benefit of GOSAT X_{CO2} data to regional CO_2 flux estimation). On an annual basis, regional uncertainty reductions over land ranged from 2% to 44%. Those reductions were shown to be variable depending on the availability of GOSAT data, which is closely related to the change of season (shifts in local solar zenith angles) and local clear sky conditions that influence success in the X_{CO2} retrieval.

Not only by changes due to season and sounding conditions, differences in X_{CO2} retrieval algorithms were also found to affect X_{CO2} spatial distributions and thus influence regional flux estimates. On a global scale, five annual total terrestrial fluxes, estimated independently from X_{CO2} datasets by five different X_{CO2} retrieval algorithms, were found to be all smaller than that estimated from the surface-only data. The spread (SD) among the five global total estimates was also found to be small. On annual regional scales, however, fluxes varied largely, particularly those estimated for the temperate Asia regions where the spatial coverage by the five X_{CO2} datasets was found to differ from one to another. In other regions where the data distribution is similar, the five flux estimates agreed well regardless of the X_{CO2} datasets used.

The influence of data spatial distribution differences on regional fluxes was further explored in the flux estimation from individual, "single-shot" X_{CO2} retrievals, as it was implied that the impact can be more pronounced without the X_{CO2} grid aggregation and averaging in inversion. Region 32, a Temperate Asia region that cover eastern China, Mongolia, the Korean Peninsula, and Japan, was chosen for the study. It was shown that five collocated X_{CO2} values found within the region and around the globe agree well; the global mean of collocated X_{CO2} SDs was 0.8 ppm, which is less than a half of X_{CO2} random error (2 ppm). The analysis of five independent flux estimates using resolution kernels indicated that constraints on fluxes are dependent on the number of "influential" X_{CO2} retrievals whose responses to the regional unit pulse emission are strong. Further, analyzing the response functions for the concerned region revealed that where in the region and how densely the X_{CO2} retrievals are distributed impact the monthly flux estimate.

The study described in Chapter 4 was performed as part of the GOSAT inversion inter-comparison campaign [Houweling et al., in review]. Through this study, the range of possible spread in regional flux estimates owing to differences in X_{CO2} retrievals was quantified. The next step to be pursued in the flux comparison effort is to quantify flux spread due to differences in existing inverse modeling systems. This evaluation is necessary in determining whether uncertainties in GOSAT-based flux estimates come from the differences in X_{CO2} retrieval algorithms or in the inversion systems themselves. Also, it allows for specifying regions in the globe where the existing inversion systems and X_{CO2} retrievals are most or least capable of estimating quality fluxes. This experiment can be performed with the use of a common set of X_{CO2} retrieval data and a priori flux data. Plans for the next-step study are being arranged.

Another potential research activity that can be conducted, upon the completion of the above-mentioned flux inter-comparison, is to find out such regions for which the inverse modeling system developed in this study is capable of generating reliable flux estimates, and to analyze their long-term trends to see if those GOSAT-optimized fluxes are in tune with observable changes including temperature, precipitation, and land cover. Findings in researching whether those trends can be explained by process-based terrestrial biosphere models can be an valuable input to ongoing studies in comparing top-down and bottom-up CO_2 flux estimates, such as one conducted by Kondo, Ichii, and Takagi [in review], and can be also a "first-step" contribution to improving the scenarios used for the prediction of future climate, as touched in the introduction in Chapter 1.

References

Houweling, S., et al., An encompassing of inverse models for estimating sources and sinks of CO_2 using GOSAT measurements, *J. Geophys. Res*., in review.

Kondo, M., K. Ichii, and H. Takagi, Comparison of the data-driven bottom-up and top-down global terrestrial CO_2 exchanges: empirical eddy flux upscaling and GOSAT CO_2 inversion, *J. Geophys. Res.,* in review.

APPENDIX: Sensitivity of flux uncertainty reduction rate (UR) to uncertainty associated with forward X_{CO2} modeling

As mentioned in Section 2.2.3, the uncertainty associated with the prediction of GOSAT X_{CO2} concentrations with NIES-TM was found out, via comparison to TCCON surface-based X_{CO2} measurements, to be 0.2% (~1 ppm) [Belikov et al., 2013]. This forward modeling uncertainty is taken into account in determining the diagonal elements of the covariance matrix $\mathbf{C_D}$; the minimum of each of the diagonal elements was set as the sum of the random error associated with GOSAT X_{CO2} (2 ppm) and the uncertainty associated with the X_{CO2} forward modeling (1 ppm). To show that the result and conclusion regarding the uncertainty reduction attained by the addition of GOSAT X_{CO2} data to the surface-based GV data are robust, I performed a check on the sensitivity of URs to the X_{CO2} forward modeling uncertainty. Here I considered a case in which the uncertainty is reduced by 50% (0.5 ppm; equivalent to doubling the current modeling capability).

Figure A1 shows annual mean URs attained with the reduced forward modeling uncertainty and they are contrasted with those presented in Chapter 3 (Figure 3.5). The average of the terrestrial URs for the 50% reduction case turned out to be 12%, which only differed by 2 points from the Chapter 3 case being contrasted. The maximum annual mean UR is found in Region 29 (47%; Temperate Asia SW: Arabian Peninsula); this region was found to be associated with the largest annual UR in the Chapter 3 case (41%). The left panel in the figure show that the high URs were attained in the estimation of fluxes for regions that are undersampled by the GV monitoring stations but well sampled by GOSAT, which is consistent with the finding presented in Chapter 3. Overall, the general conclusion of Chapter 3 is found to be not affected significantly by changes in the performance of the atmospheric tracer transport model used.

Additionally, I also considered a possible future case in which the random error associated with the current versions of X_{CO2} retrieval datasets (2 ppm; Sections 4.2.1 and 5.2.1) is reduced by 50% through improvements in the retrieval algorithms (1 ppm). Figure A2 shows the annual mean URs obtained with the 50%-reduced X_{CO2} random error; the forward modeling uncertainty in this case was kept to 1 ppm. The average of the terrestrial URs for this case was 15%, and the maximum annual mean UR is found again in Region 29 and it exceeds 50% (53%). The overall high-low patterns of the UR distribution remains nearly the same as the above-mentioned forward modeling 50% error-cut case, but now many of the undersampled regions attain URs greater than 20-25% (more toward blue color). The results shown suggests that the 50% X_{CO2} random error cut may lead to attaining greater annual URs than the 50% forward modeling uncertainty cut.

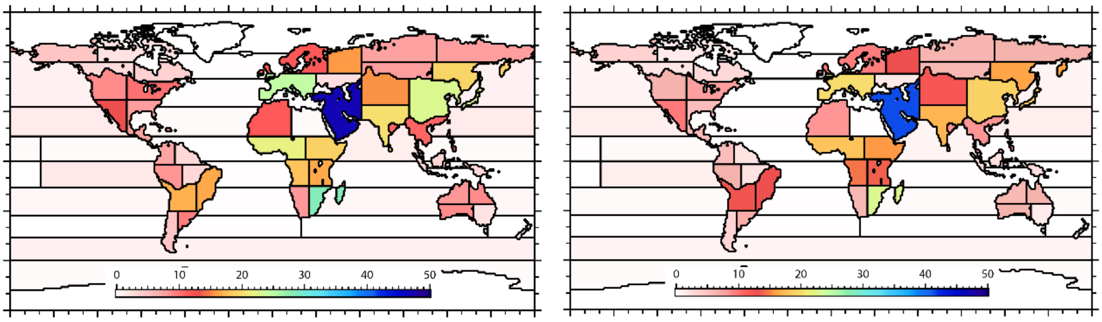

Figure A1. Left: annual mean UR attained with forward X_{CO2} modeling uncertainty reduced by 50% (0.5 ppm). Right: the same as Figure 3.5 (Chapter 3). Values shown are in %.

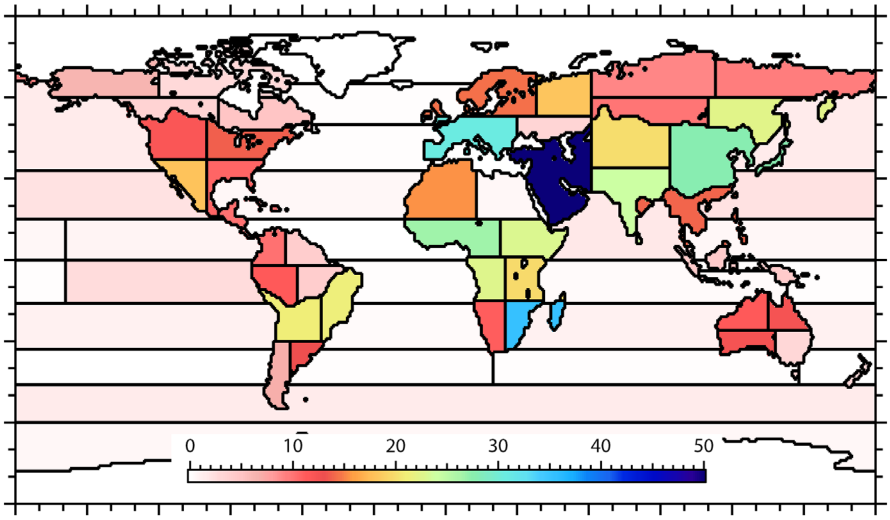

Figure A2. Annual mean UR attained with X_{CO2} random error reduced by 50% (1 ppm).

Values shown are in %.