The Pennsylvania State University

The Graduate School

Department of Meteorology

REDUCING THE UNCERTAINTY OF NORTH AMERICAN CARBON FLUX ESTIMATES USING AN EXTENDED ATMOSPHERIC CARBON DIOXIDE MEASUREMENT NETWORK

A Dissertation in

Meteorology

by

Martha Peirce Butler

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2010

The dissertation of Martha Peirce Butler was reviewed and approved* by the following:

Kenneth J. Davis
Professor of Meteorology
Dissertation Advisor
Chair of Committee

Raymond Najjar
Associate Professor of Oceanography

Sukyoung Lee
Professor of Meteorology

Klaus Keller
Associate Professor of Geosciences

A. Scott Denning
Professor of Meteorology
Colorado State University

William H. Brune
Distinguished Professor of Meteorology
Head of the Department of Meteorology

*Signatures are on file in the Graduate School

ABSTRACT

We evaluate North American carbon fluxes using a monthly global Bayesian synthesis inversion that includes well-calibrated carbon dioxide concentrations measured at continental flux towers. We employ the NASA Parameterized Chemistry Tracer Model (PCTM) for atmospheric transport and a modified version of the inversion used by the Atmospheric Tracer Transport Model Intercomparison Project (TransCom) with sub-continental resolution and annual variability of transport. We sub-sample carbon dioxide time series at four North American flux tower sites for mid-day hours to ensure sampling of a deep, well-mixed atmospheric boundary layer. The addition of these flux tower sites to a global network reduces North America mean annual flux uncertainty for 2001-2003 by 15% to 0.4 Pg yr^{-1} compared to a network without the tower sites. North American flux is estimated to be a net sink of 1.3 Pg C yr^{-1}, within the uncertainty bounds of the result without the towers. Uncertainty reduction is found to be local to the regions within North America where the flux towers are located. Including the towers reduces covariances between regions within North America. We estimated potential future uncertainty reduction with simulated observations at North American sites that are now or planned to be instrumented for suitable carbon dioxide measurements.

We also tested a micrometeorological adjustment to surface carbon dioxide measurements to approximate mid-continental-boundary-layer measurements. This adjustment can be calculated during mid-day hours using atmospheric measurements commonly available at flux tower sites. The atmospheric transport models used in global atmospheric inversions often do not have sufficient spatial or temporal resolution to capture small-scale variability in the continental boundary layer. We find that using mid-day hours of observations at continental sites, either with or without the micrometeorological adjustment, allows inclusion of continental sites in global atmospheric inversions. Increased continental observation density is necessary for estimating carbon fluxes with finer resolution in space and time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Portions of this dissertation are based on multi-authored manuscripts that have been or will be submitted for publication. Chapter 2 is a modified version of a manuscript submitted to the International Meteorological Institute in Stockholm and Blackwell Publishing for publication in the journal *Tellus Series B: Chemical and Physical Meteorology* special issue for the International Carbon Dioxide Conference (ICDC8). Material from Chapters 2 and 5 was presented at this conference. Martha Butler, the author of this dissertation, is the first author of all of this work. She was primarily responsible for developing the experiments, acquiring the data, and conducting the data processing, analysis, writing and revisions. Co-authors named in the footnote for Chapter 2 and committee members were involved in methodological guidance, editing and funding.

ACKNOWLEDGEMENTS

Undertaking this work has been an adventure and a learning experience. It would not have been possible without the encouragement of many, especially my adviser Ken Davis. Support from the rest of my committee, Sukyoung Lee, Ray Najjar, Klaus Keller, and Scott Denning is truly appreciated. Past and present members of the Davis and Denning research groups have provided advice and assistance. Visits to Colorado State and those introductory discussions about atmospheric inversions with Scott Denning and Kevin Gurney served to shape what this work has become. Randy Kawa, technical adviser for my NASA GSRP fellowship, continues to treat me as part of his extended research group, educating me in more of the global carbon cycle than I ever imagined. The faculty of Penn State's Department of Meteorology deserves special thanks for taking a chance in admitting me to the graduate program 30 years after my undergraduate degree, and then supporting me to make sure their initial decision was correct. I must also thank family and friends, who did not laugh too hard when I said I was finally ready to go back to school. You have all been wonderful.

I have enjoyed my participation in and contact with the Atmospheric Tracer Transport Model Intercomparison Project (TransCom) modeling community. Their insight, drawn from long experience, has been invaluable. They also made available the inversion code used in the TransCom Interannual Variability experiment, giving me a solid start on my own version of the code. In return, I have contributed results of this experiment to the TransCom community and to the North American Carbon Program Interim Synthesis activity.

This work draws on the efforts of agencies world-wide that sponsor carbon dioxide measurement programs and cross-network calibration activities. Their belief that the data should

be disseminated and used by the scientific community at large is commendable. Special thanks to the NOAA Earth System Research Laboratory for their leadership in this effort, and to the World Data Center for Greenhouse Gases and the Ameriflux network for making data available.

Financial support for this work came from the following: NASA Graduate Student Research Program fellowship (technical adviser: S. R. Kawa at the NASA Goddard Space Flight Center); PSU College of Earth and Mineral Sciences (Carl H. and Helen H. Chelius Graduate Fellowship in Earth Science and Centennial Graduate Research Award); NSF Research Collaboration Network (RCN) Grant No. DEB-0130380 enabling research travel support for visits to the Denning group at Colorado State University; U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA), Office of Global Programs, Global Carbon Cycle program Grant No. NA040AR4310124; Office of Science (BER), U.S. Department of Energy. Any opinions, findings, and conclusions or recommendations herein are those of the author and do not necessarily reflect the views of the cooperating agencies.

Chapter 1

Introduction

1.1. Background

This work is a contribution to the ongoing study of the contemporary global carbon cycle from the perspective of the atmospheric carbon dioxide budget. About half of the anthropogenic carbon emitted through fossil fuel burning, cement manufacturing and land use change each year remains in the atmosphere. The other half is being incorporated into natural exchanges of carbon between the atmosphere and terrestrial vegetation and the ocean surface. Questions remain about the spatial and temporal resolution of these exchanges and, therefore, about the specific mechanisms involved. Transport and mixing in the atmosphere, at short time scales compared to geological mechanisms, further confound the problem. Understanding the causes of the annual variability of the exchanges is an additional goal.

Decades of precise measurements of carbon dioxide in the atmospheric boundary layer and ocean surface layer have made clear that there is a latitudinal gradient in carbon dioxide that points to a carbon sink in the Northern Hemisphere as a missing term in the atmospheric carbon budget. Tans et al. (1990) showed that the north-south gradient of carbon dioxide concentration and the best available air-sea flux estimates dictated that this sink is terrestrial and in the Northern Hemisphere. Different approaches and different time periods of study yield little agreement as to whether this terrestrial sink may be in North America or Eurasia. We do know that the annual variability of the exchanges is dominated by terrestrial fluxes, and yet the measurements of carbon dioxide have been primarily sampling marine boundary layer air. Incorporating

measurements from the continental boundary layer into the diagnosis of the spatial and temporal resolution of these exchanges is the focus of this work.

The history of the development of the global atmospheric inversion as an approach to the problem of determining global carbon sources and sinks is documented in Chapter 14 of Enting (2002). Earliest attempts to tackle the problem from the atmospheric perspective began with basic mass balance methods using box models. The inversion method used in the experiment described in this work evolved as computing resources have grown, making it possible to use simulated transport in place of simple exchange rates between atmospheric boxes. For an idealistic view of the synthesis inversion method used in this work, imagine a gridded atmospheric model with perfect transport. Assign a unit amount of constant carbon dioxide emission to each surface layer grid cell to model the exchange between atmosphere and land and atmosphere and ocean. Run the transport model long enough to establish a stable spatial pattern of the carbon dioxide in the whole model atmosphere. Assume that there is a precise measurement of carbon dioxide in real space corresponding to every grid cell in the model atmosphere. Compare the transport model predictions in each grid cell to the precise measurements. An inverse method is used to find the adjustments to the unit surface grid cell emissions to make the model predicted data match the measurements. In reality, we do not have optimally distributed precise measurements, and the transport models are not perfect. We have a mixed-determined problem (Menke, 1984), where there may be more than one solution that fits the available measurements. For example, a measurement of the carbon dioxide mixing ratio could be consistent with either small local emissions or large emissions at a distance that have been transported and mixed throughout the atmosphere. The spatial distribution of measurements is not of high enough density to determine which of these is correct.

1.2. Assumptions and Definition of the Problem Space

To make the problem tractable, two strategies are typically used: reduce the number of model parameters and add prior information as a first guess to the eventual solution. Both strategies are used in this work. The most obvious reduction in model parameters is spatial reduction, solving for fluxes in large regions. This strategy has been used in similar research where the regions are defined at continent and ocean basin scale (e.g., Gurney et al., 2002). As others have done (e.g., Patra and Matsyukov, 2003), we divide most of these large regions to test whether the increased number of observation sites we use is enough to constrain smaller regions. Earlier inversions also solved for annual mean fluxes over periods of five years or more (Gurney et al., 2002) or a seasonal mean representative of a climatologic year (Gurney et al., 2004). In this work we solve for each region and month (e.g., Baker et al., 2006) over a period of five years, not enough to establish long term trends, but enough to show some degree of annual variability.

Although the global carbon dioxide observation network is growing, it is still sparse for use in global atmospheric inversions. Efficient atmospheric mixing and spatial gaps in the observation network conspire to make the problem ill-defined. We use existing estimates of anthropogenic emissions from fossil fuel and biomass burning and cement manufacturing, as well as natural terrestrial and ocean fluxes as a starting point. These are referred to as background fluxes in this work. The ocean and terrestrial exchange background fluxes are adjusted within the constraints of assumed uncertainties in these background exchanges. Fossil fuel and biomass burning emissions are assumed to be well-defined and are accepted as truth. If they are not correct, then the adjustments to the terrestrial and ocean fluxes will include the changes to these fixed fluxes; the inversion-determined adjustments are made to the overall net fluxes.

The changes to the methodology used in this project are evolutionary, not revolutionary; we address some, but not all of the assumptions made in previous work. These assumptions and

their impact on our understanding of the inversion results are discussed in Chapter 3 and in the summary in Chapter 6. Other methods have also used to attempt to achieve the same goals. Each of these methods is subject to its own limitations. The Bayesian synthesis inversion method that we use can still be useful as, for examples, a first assessment of incorporation of column averaged carbon dioxide measurements from space or to determine boundary conditions for regional atmospheric inversions. The method can also be expanded, as others have done (Rayner et al.,1999; Peylin et al., 2001) to incorporate measurements of $^{13}CO_2$ and O_2/N_2 in an effort to distinguish between land and ocean fluxes.

1.3. Dissertation Outline

Chapter 2 contains the description of the global atmospheric inversion method and the impact on the North American flux uncertainty of including continental, quasi-continuous observations of carbon dioxide. The control inversion, the continental extension network inversion in Chapter 2, is the basis for the sensitivity tests of the method in Chapter 3. Chapter 4 covers issues of sensitivity to network design choices. Chapter 5 introduces a micrometeorological adjustment to continental surface measurements of carbon dioxide to approximate mid-boundary-layer values, the virtual tall tower (VTT) adjustment. Results are presented for a variation of the control inversion in which VTT-adjusted observations at North American flux towers are used in place of the surface observations at these locations. The impact of future North American observation sites is also explored in Chapter 5 with inversions using simulated observations at locations currently instrumented for carbon dioxide measurements calibrated to global standards. Chapter 6 is a summary of the findings and recommendations for future network design and use of the global atmospheric inversion method. Assumptions are also reassessed in light of the findings. The Appendix contains material that is common to the entire

work including data sources, observation site details and citations, and composition of the networks of observing sites used in the inversion variations.

Chapter 2

Impact of Continental, Quasi-Continuous Observations of Carbon Dioxide in the Estimation of the North American Carbon Sources and Sinks[1]

2.1. Introduction

About half of the anthropogenic carbon emitted into the atmosphere remains in the atmosphere each year. The remainder is taken up by the ocean and terrestrial ecosystems through the processes responsible for the natural exchange of carbon between the atmosphere and terrestrial vegetation and the surface ocean (Forster et al., 2007; Denman et al., 2007). Numerous studies (Myneni et al., 1999; Nemani et al., 2003; Potter et al., 2003) show that climate cycles, local weather, and ecosystem conditions all affect the interannual variability of this uptake of carbon. Our understanding of the mechanisms governing the dynamics of the carbon cycle has been hampered by a limited ability to locate and quantify these exchanges at sufficiently fine temporal and spatial resolution (Bousquet et al., 2000; Gurney et al., 2002; Ciais et al., 2005; Baker et al., 2006; Peters et al., 2007). Accurate and precise quantification of sources and sinks at regional and continental scales is likely to be increasingly important for evaluation and monitoring of carbon management policies.

Global atmospheric inversions have been used to infer sources and sinks of carbon (both natural and anthropogenic) at continental and ocean basin scales from atmospheric measurements of carbon dioxide using tracer transport models. Model intercomparison projects, including the TransCom (Atmospheric Tracer Transport Model Intercomparison Project) series (Gurney et al.,

[1] This chapter is a modified version of a manuscript submitted to the International Meteorological Institute in Stockholm and Blackwell Publishing for publication in *Tellus B: Chemical and Physical Meteorology*: Butler, M.P., Davis, K.J., Denning, A.S. and Kawa, S.R. Using continental observations in global atmospheric inversions of CO_2: 1. North American carbon sources and sinks.

2002; Gurney et al., 2004; Baker et al. 2006), have been designed to attribute the uncertainties in the continental and ocean basin fluxes estimated by this method. These studies show that transport model differences and the uneven and sparse global distribution of atmospheric carbon dioxide measurements both contribute to the uncertainty of the inverse flux estimates. While transport models are improving and the global measurement network for carbon dioxide is expanding, there is still a fundamental representation error (Kaminski et al., 2001; Engelen et al., 2002) inherent in the global atmospheric inversion method. There is a mismatch in space and time resolution between the transport models (grid boxes and minutes), the observations (points in space and time), and the inversion solution (continents or sub-continents and months or weeks). Observations, subject to local atmospheric variations, and transport models, for which the mesoscale atmospheric variations are sub-grid scale, constrain continental and ocean basin results. If the inversion solution is at the continental scale, the continental results cannot be effectively separated by sub-regions within the continent. The strengths of this continental-scale global inversion method are the minimum number of unknowns and the global coverage. The strong and uncertain assumptions about correlations of fluxes in space and time are a weakness. The assumptions, however, are essential; surface observations will always be uneven and sparse.

The ideal situation would be to invert on the grid and time scale of the transport model (Engelen et al., 2002). Global atmospheric inversions at the spatial resolution of the transport model (e.g., Rödenbeck et al., 2003a; Peters et al., 2007; Mueller et al., 2008; Gourdji et al., 2008) aim for the finest resolution possible to minimize representation assumptions. Regional atmospheric inversions target a geographically limited domain with finer spatial and temporal resolution (Gerbig et al., 2003; Peylin et al., 2005; Lauvaux et al., 2008; Schuh et al., 2009). Both of these approaches involve many more unknowns, which cannot be resolved independently given the current observation density. The underlying assumptions may be minimized, but at the expense of complexity and computational costs.

In this experiment we take a pragmatic, middle-ground approach to the big-region global inversion by choosing a number of regions roughly matched to the observation density currently available. We expect that posterior uncertainties and spatial correlations will be reduced, and that the problem will be computationally tractable using simple inversion methods. Inversion results can be aggregated to the larger TransCom continental regions for comparison with published results. We can also test the ability of the expanded network to constrain the smaller regions with this method.

Typically, global atmospheric measurement network sites have been chosen to facilitate sampling background concentrations of trace gases including carbon dioxide. These background measurement networks have yielded important understanding of interhemispheric gradients in carbon dioxide mixing ratios (Tans et al., 1990; Denning et al., 1995; Keeling et al., 1996) and of the mean annual cycles of carbon emissions and uptake (e.g., Keeling et al., 1995). These data, however, provide limited understanding of the continental carbon cycle. We cannot diagnose continental or regional scale fluxes and determine factors influencing terrestrial fluxes without observing sites over the continents. Continental carbon dioxide measurements are characterized by strong diurnal and seasonal cycles that reflect a combination of biological fluxes and atmospheric boundary layer dynamics (Bakwin et al., 1998; Yi et al., 2001; Davis et al., 2003). Continental data also contain strong gradients driven by weather (e.g., Hurwitz et al., 2004; Wang et al., 2007; Parazoo et al., 2008). These strong, rapidly varying gradients may be difficult to simulate, but the continental data contain information needed to resolve regional sources and sinks of carbon with increasing spatial and temporal resolution.

In this paper we use the Bayesian synthesis inversion method to demonstrate the impact of including more continental measurement sites in the global measurement network. The added sites are long-running eddy covariance flux towers with high precision carbon dioxide measurements calibrated to global standards. Carbon dioxide measurements at flux towers do not

need to be calibrated to global standards for the calculation of net ecosystem exchange of carbon dioxide using the eddy covariance method. The sites used in this study, however, are part of a growing network where the calibration to global standards is done with the intent of providing data suitable for application to atmospheric inversion studies. The five towers chosen have data available during the 2000-2004 time period of this study. The potential effect of adding increasing numbers of well-calibrated flux towers will be explored in future research.

We focus here on the effect on the North American carbon balance, recognizing the danger that, in an ill-conditioned problem such as this, increasing the density of observations in North America can introduce dipoles (attributing fluxes to North America based on the observations, and offsetting these with opposing fluxes in less well-constrained regions). With the exception of Boreal Asia, the Northern Hemisphere is well represented in the networks tested. Concentrating observation sites in North America and Europe may result in compensating fluxes in Boreal Asia and northern ocean regions. Published inversion results do not necessarily agree on the partitioning between Northern Hemisphere land regions, nor do they agree with bottom-up estimates of carbon fluxes from biogeochemical models. Using the example of the balance between Europe and extratropical Asia land sinks found in recent inversions, some results find a larger sink in Europe (Baker et al., 2006; Mueller et al., 2008) and others in Asia (Peters et al., 2007; Rödenbeck et al., 2003a; Rödenbeck et al., 2003b). We will examine our results in this light, but concentrate on the uncertainty improvement of the added measurement sites in this paper.

In Section 2.2 we describe the estimation method. In Section 2.3 we present global and North American results for two typical global measurement networks and a third network including five additional continental sites. Section 2.4 includes a discussion of issues. We conclude in Section 2.5 with recommendations for applicability of the method to future experiments.

2.2. Methods

2.2.1. Estimation Method

The Bayesian synthesis inversion method used is shown in **Figure 2.1** and described by Enting (2002) and Tarantola (2005), with the experimental protocol (Gurney et al., 2000) following closely that of the TransCom interannual variability model intercomparisons (Baker et al., 2006; Gurney et al., 2008). We depart from the TransCom inversion method in a few important respects. 1) For observation data, we use monthly means and standard deviations derived directly from site observations of carbon dioxide instead of using a smoothed data product and imposed uncertainty. 2) We use interannually varying meteorological driver data in our transport modeling. 3) Finally, we include biomass burning emissions explicitly.

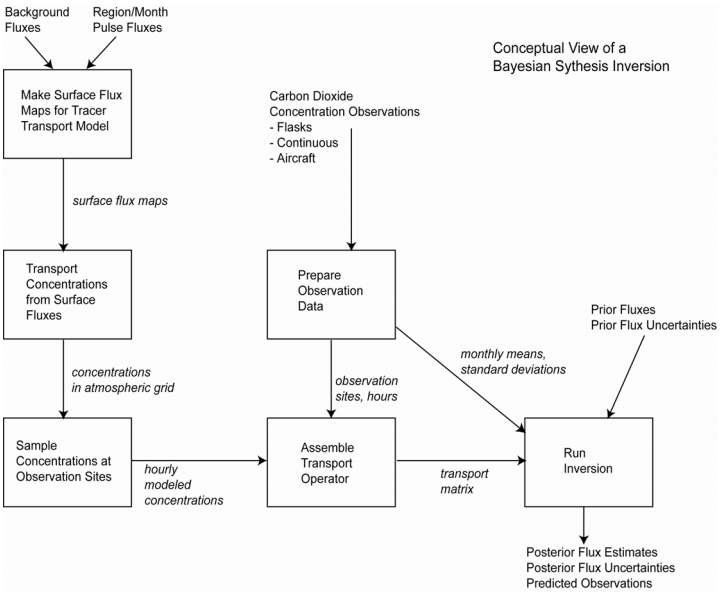

Figure **2-1**: Overview of the Bayesian synthesis inversion method

The solution is for monthly carbon source/sink estimates (2000-2004) for 47 sub-continental regions and ocean basins using monthly mean carbon dioxide mixing ratio measurements. **Figure A-1** shows the region definitions and **Figure 2-2** the locations of the observing sites. The problem is ill-constrained due to the sparse and uneven distribution of observations; the method solves for adjustments to natural land- and ocean-atmosphere exchanges (referred to here as background fluxes) within the constraint of imposed prior uncertainties. Fossil fuel emissions and biomass burning emissions are assumed to be correct and not adjusted in the inversion process.

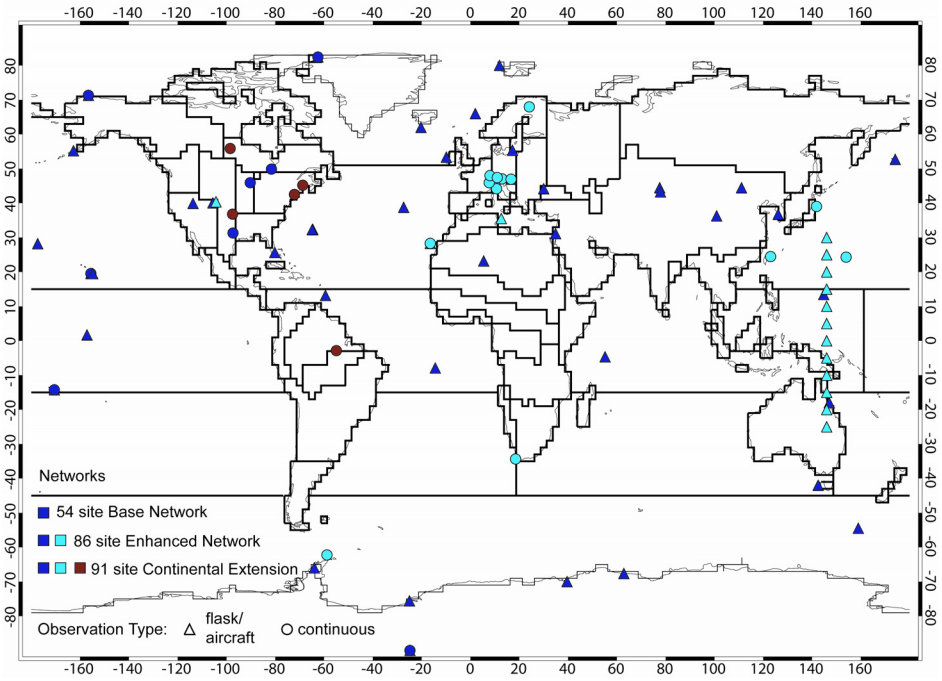

Figure **2-2**: Observation sites used in the three networks tested. Symbol colors denote network composition: base network (blue); enhanced network (blue and cyan); continental extension network (blue, cyan, and dark red). Symbol shapes denote the type of observation: quasi-continuous (circles) and discrete (triangle).

Following Baker et al. (2006), the atmospheric carbon dioxide at a measuring site can be represented as the linear combination of the responses at the location to the background fluxes and to the unknown adjustment fluxes from each of the regions and months,

$$\mathbf{c_{obs}} = \mathbf{c_{fwd}} + \mathbf{Hx} \tag{2.1}$$

where $\mathbf{c_{obs}}$ is the time series of monthly carbon dioxide observations (length of vector is number of observation sites x 60 months), $\mathbf{c_{fwd}}$ is the modeled concentration time series using the background fluxes, $\mathbf{H}$ is a transport matrix (described below), and $\mathbf{x}$ are the unknown monthly adjustments to the background terrestrial and ocean fluxes (length of vector is 47 regions x 60 months).

Solving for the unknown monthly adjustments $\mathbf{x}$ is done using a singular value decomposition approach (Rayner et al., 1999) to the minimization of the cost function

$$J = (\mathbf{c_{obs}} - \mathbf{c_{fwd}} - \mathbf{Hx})^T \mathbf{R}^{-1} (\mathbf{c_{obs}} - \mathbf{c_{fwd}} - \mathbf{Hx}) + (\mathbf{x} - \mathbf{x_0})^T \mathbf{P}_0^{-1} (\mathbf{x} - \mathbf{x_0}) \tag{2.2}$$

where $\mathbf{R}$ is the covariance matrix specifying the observation, transport and representation error, $\mathbf{x_0}$ are *a priori* estimates of the solution, and $\mathbf{P_0}$ is the covariance matrix of uncertainties of these *a priori* estimates. In this experiment both $\mathbf{R}$ and $\mathbf{P_0}$ are described by diagonal matrices, as has been common practice. This cost function minimization is a least squares solution weighted by the variability of the observations and penalized for deviation from the *a priori* estimates. The analytical solution for the flux adjustments $\hat{\mathbf{x}}$ is

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}_0^{-1}\right)^{-1} \left(\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{c_{obs}} - \mathbf{c_{fwd}}) + \mathbf{P}_0^{-1} \mathbf{x_0}\right) \tag{2.3}$$

and the *a posteriori* covariance matrix of flux errors $\mathbf{P}$ is isolated from

$$\mathbf{P}^{-1} = \mathbf{P}_0^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \tag{2.4}$$

allowing explicit examination of the covariances for independence of the solution. This analytical posterior uncertainty is a function of the uncertainty of the prior flux estimate and the uncertainty

attributed to the observations and assumes that the method is appropriate to the problem. As a consequence, the accuracy of the posterior covariance matrix is a function of the accuracy of assumptions inherent in the *a priori* uncertainty and data uncertainty assignments. Another product of the inversion method is a set of predicted observations, the carbon dioxide concentration values that would be expected at each observation site given the *a posteriori* flux solution. The predicted observations provide a basis for analysis of the data residuals, which cannot be done with the fluxes.

The analytical solution delivers an adjustment to the background fluxes for each region and month in 2000 through 2004. Results for the central years 2001-2003 are retained for analysis. The first and final year of the solution are discarded to minimize edge effects, including inaccuracies introduced by the spinup method described below and the length of time it takes an atmospheric signal to reach distant observation sites. All results reported here are the adjusted land and ocean fluxes including the assumed biomass burning emissions for each month in order to compare with published results. The assumed background fossil emissions are not included. Choices for the components of the method shown in **Figure 2-1** follow.

2.2.2. Background Fluxes

The terrestrial background flux is the interannually varying hourly SiB3 terrestrial flux (Baker et al., 2007) as prepared for the TransCom continuous experiment (Law et al., 2008; Patra et al., 2008). An alternative terrestrial flux, the CASA monthly climatology (Randerson et al., 1997) used in previous TransCom experiments, is also tested to investigate the solution dependence on the terrestrial background flux used and the time resolution of the flux. The ocean background flux is the monthly climatology of Takahashi et al. (2002), also as used in the TransCom continuous experiment. Two fossil fuel emissions fluxes are tested: the seasonally

varying emissions described by Erickson et al. (2008) and a non-seasonally varying version. Both are based on the 1995 spatial distribution of Brenkert et al. (1998) and Li et al. (1996), with annual totals scaled to the appropriate years based on Marland et al. (2007). The interannually-varying monthly biomass burning emissions are from the Global Fire Emissions Database version 2 (GFED2) (Giglio et al., 2006; van der Werf et al., 2006). Each background flux (4 tracers x 5 years) is regridded to the horizontal spatial grid of the transport model and used as a surface boundary condition for a forward run from the beginning of the applicable year to the end of 2004. The model output is sampled hourly to allow for matching of sampling between model and observations (see section 2.2.8.) when constructing the $\mathbf{c}_{\mathbf{fwd}}$ time series represented in Equation 2.1.

2.2.3. Transport Response Functions

The 47 regions specified for the solution are shown in the Appendix in **Figure A-1** and **Table A-1**. The 11 ocean regions are the same as those used in the TransCom experiments. The 36 land regions are conformable to the TransCom land regions. The Boreal Asia, Temperate Asia, Tropical Asia, and Australia regions are defined as in TransCom. Remaining TransCom regions are sub-divided into sub-regions: Boreal North America (3 regions); Temperate North America (7); Tropical America (3), Temperate South America (2), Europe (7), Northern Africa (5) and Southern Africa (5). These continental regions are sub-divided to test the size of the regions for which flux solutions can be constrained given current and future observing networks. The results for these regions are aggregated to larger regions for reporting and comparison with published results.

To construct the transport matrix $\mathbf{H}$, emissions of known magnitude (one month duration at an annual rate of 1 Pg carbon emission) are run forward as inert species through the transport

model for each region-month (47 regions x 12 months x 5 years), beginning in the applicable month and year and ending after 25 months of transport or the end of 2004. Model output is sampled hourly at each observation location through the 25 months (or less) of each region-month transport run, and then assumed to remain at a constant well-mixed, residual level after the transport is terminated through the end of 2004. Longer transport runs showed that the assumption of a well-mixed concentration beyond 25 months is justified.

The flux pattern within each terrestrial region is based on annual NPP simulated by CASA monthly climatology (Randerson et al., 1997), scaled to the transport model grid cells within each region. The pattern is derived from the similar regional patterns in the TransCom experiments (Gurney et al., 2000) and regridded to the spatial resolution of the transport model used here. There is no variation by model grid cell within the ocean regions. These flux patterns constitute a hard constraint in this inversion method. The inversion solution specifies the adjustment to the background flux for the region as a whole; the distribution of the flux within the region is fixed by this flux pattern.

2.2.4. Tracer Transport Model

The tracer transport model used in this experiment is the Colorado State University version of the NASA Parameterized Chemistry Tracer Model (PCTM), described in Kawa et al. (2004) which has participated in TransCom experiments. Winds, temperatures, diffusion coefficients, and convective mass fluxes are from the NASA Goddard Earth Observation System 4 (GEOS-4) data assimilation system (Bloom et al., 2005). The 6-hourly meteorological driver data are linearly interpolated to the 15-minute time step of the model. The model is run on a 2.5° longitude by 2.0° latitude grid, with 25 hybrid vertical layers. The model exhibited intermediate performance in the TransCom Interannual Variability model intercomparisons (Baker et al., 2006;

Gurney et al., 2008) and is also one of the three transport models used in the Gurney et al. (2005) study of the potential bias in inversions caused by using non-varying background fossil emissions. PCTM was chosen for its intermediate performance among the TransCom models with relatively strong trapping of non-varying fossil fuel emissions but a relatively weak rectifier for the terrestrial background flux due to seasonal differences in the vertical mixing close to the model surface. Additional detail about the transport model performance can be found in Parazoo et al. (2008).

To establish the atmospheric interhemispheric gradient of carbon dioxide, we used a spin up procedure that differs from the technique used in TransCom-IAV (Baker et al., 2006). In TransCom –IAV, the same 36 months of response to the background fluxes (with appropriate annual scaling for the fossil fuel emissions) are reused each year in the construction of the transport matrix $\mathbf{H}$. The first few years in the transport matrix do not contain residual responses from earlier years; standard practice is to discard the first few years of the inversion solution before analysis of the results. In our case, we have added residual second, third, and fourth year responses to the background fluxes to the first year (2000) responses to represent cumulative responses from the prior years (1999, 1998, and 1997). We then discard this first year of the inversion results before analysis to minimize any potential effects of this spinup method. There will be inconsistencies to the extent of the differences between the proxy and actual features of meteorological driver data and annually varying spatial patterns of the terrestrial and biomass burning background fluxes.

2.2.5. A Priori Constraints

The *a priori* constraint consists of a source/sink term and an uncertainty specification for each region-month flux adjustment in the inversion solution and for the background fluxes. The

region-month *a priori* adjustments in this experiment are set at zero, implying "no correction" to the fossil fuel and biomass burning background emissions and no additional land-use change terrestrial flux. The terrestrial background flux is annually neutral; the ocean background flux specifies a global annual sink of 1.6 Pg C; biomass burning emissions are ~2 Pg C annually; fossil flux emissions are ~7 Pg C annually for this period. In order to balance the global growth in atmospheric carbon dioxide concentration, there is an expectation that the inversion-adjusted terrestrial and ocean region-month fluxes will be a net sink. The region-month uncertainties (diagonal values in the $\mathbf{P}^0$ prior flux covariance matrix) vary by month, summing globally to 5.4 Pg C annually. This range is larger than the 2.76 Pg C annual uncertainty used in the TransCom interannual variability control (Baker et al., 2006) and network sensitivity (Gurney et al., 2008) experiments. The magnitude of the monthly uncertainties is intended to be loose enough to allow the inversion to make substantial adjustments to the background terrestrial and ocean fluxes while remaining biogeochemically realistic. These region-month uncertainties are calculated as the sum of the magnitudes of three months of the applicable background flux (terrestrial or ocean) for the region centered on each month. This allows the inversion "room" to correct for timing differences in seasonal cycles and generally allows for more latitude for adjustment in months when the fluxes are of largest magnitude.

2.2.6. Observation Networks

Observation site locations are shown on the map in **Figure 2-2**. A list of the agencies supporting carbon dioxide measurement programs is in the Appendix in **Table A-2**. Complete lists of all observation sites including site names, locations, responsible agencies, and ranges of monthly variability of the observations can be also be found in the Appendix in **Table A-3** (quasi-continuous measurement programs), **Table A-4** (discrete surface measurement programs) and

Table A-5 (discrete aircraft measurement programs). A cross-reference of observation sites by site code and network composition for the networks Chapter 2 is given in **Table A-8**. The observation sites include observatories, tall towers, flask sampling sites, and an aircraft vertical profile from the NOAA ESRL network (//www.esrl.noaa.gov/gmd/ccgg/index.html), sites from other measurement programs with data archived at the World Meteorological Organization (WMO) Global Atmosphere Watch World Data Center for Greenhouse Gases (//gaw.kishou.go.jp/wdcgg/wdcgg.html), and carbon dioxide time series from five flux towers. The goal is to use measurements calibrated to the WMO standards for carbon dioxide (Tans et al., 2003; Zhao et al., 1997; Zhao and Tans, 2006). Calibration to WMO standards across multiple measurement methods and agencies has proved to be a challenge (Masarie et al., 2001). Law et al. (2003) show that the impact of inter-agency calibration offsets can be accommodated in inversions on synoptic time scales (~5days), but may be significant on the monthly time scale we use in this experiment. To the extent that we have included observations from multiple measurement programs, we may have introduced bias into our results.

Stations are selected based on data availability to minimize bias introduced by gap filling. An upper limit of 12 missing months during the 2000-2004 time period was allowed, with a few exceptions for sites which began operation after 2000 (tall tower in Moody, Texas; flux towers at Southern Great Plains and Tapajos). This requirement prohibited the use of a number of sites that would be used if a smoothed, extrapolated data product such as GLOBALVIEW (GLOBALVIEW-CO2, 2007) were used. Three networks are tested in this experiment. The first is a 54-site base network, sourced from three agencies: National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA ESRL), Environment Canada (EC), and CSIRO Marine and Atmospheric Research GASLAB in Australia. All of the observation sites in this network are also used in CarbonTracker (Peters et al., 2007; CarbonTracker 2008, http: //carbontracker.noaa.gov); we cannot use some of the CarbonTracker network sites due to

the completeness of record requirement of our method. An 86-site enhanced network includes sites from other agencies, including high altitude aircraft and mountain top sites, archived at the World Data Center for Greenhouse Gases. These sites are used to improve global coverage. The last network, the 91-site continental extension network, adds the five flux tower observation sites. A primary objective of this experiment is to show the effect on the inversion estimates of including the four added North American continental sites.

2.2.7. Observation Data Preparation

Observations are selected with minimal screening for data quality issues identified by the observing agencies. For quasi-continuous continental sites, mid-day hours are selected (12-16 LST (local standard time)), except for mountain top sites where mid-night hours are used. This selection is intended to maximize the contribution to the monthly mean of the hours representing well-mixed atmospheric conditions; these are conditions most likely to be modeled correctly in the tracer transport model. Monthly means are calculated from all available hours in the daily selection time periods without regard to meteorological conditions. Data uncertainties are computed as the standard deviations of these monthly means, with an imposed minimum of 0.50 ppm. These uncertainties are used to populate the variances in the diagonal of the data covariance matrix $\mathbf{R}$ (Equations 2.1, 2.3, 2.4).

For this analytical solution method it is necessary to fill gaps in the data for months with no observations. For observation sites that are represented in GLOBALVIEW (GLOBALVIEW-CO2, 2007), gaps are filled with the average of the GLOBALVIEW site data for the weeks corresponding to the month to be filled. For stations with no GLOBALVIEW representation, a climatology is constructed of the monthly departure of existing observations at the station from the GLOBALVIEW marine boundary layer (MBL) value for the station latitude. This difference

climatology is applied to the MBL value for the missing months to fill the gap. The uncertainty for each gap-filled month is assigned as either the climatologic observation variability for the month for the station in 2000-2004 or a mean variability for all months with existing observations for the 2000-2004 time series for the station.

We have not added an additional model-data mismatch error to the data uncertainties in $\mathbf{R}$ to account for errors in model transport or the representation of regions by points (observation sites); we have assumed that the monthly variability of the observation time series accommodates the additional error. As this is not likely to be true, we will revisit this decision in Chapter 3. **Table A-3**, **Table A-4**, and **Table A-5** include the range of monthly variability for each station. Compared to the TransCom interannual variability experiment (Baker et al., 2006), these uncertainties are larger by ~0.20 ppm at the minimum and up to two times larger at the maximum for remote sites. For continental and coastal sites, the minimum uncertainties are comparable and maximum uncertainties are at least as large as those in Baker et al. (2006), which used a smoothed data product for the observation data time series and included model-data mismatch error for some observation sites.

2.2.8. Model Sampling

Model sampling is accomplished by saving vertical profiles of tracer concentration, pressure and temperature at the model grid cell nearest to each observing site hourly for the duration of the forward runs of the tracer transport model. Coastal stations are sampled in an off-shore grid cell to approximate observation protocol for flask sampling of background air. For surface observing stations the concentration samples are taken from the surface layer in the chosen grid cell. For elevated stations, vertical pressure interpolation is used to resolve the concentration sample to an equivalent elevation. Monthly mean model sample concentrations and

standard deviations are calculated from the same hours for which observations are chosen, a concept referred to as co-sampling (Peters et al., 2007). For flask and aircraft observation sites (discrete, rather than continuous, sampling), three hours of model samples centered on the hour of the observation are used to smooth discrepancies in timing of weather events in the transport. Monthly mean model samples are then calculated from these co-sampled hours from the forward runs for each of the region-month pulses and the background fluxes. The transport matrix $\mathbf{H}$ (equation 2.2) is populated with the monthly mean concentration responses at the observation sites to the region-month pulses and background fluxes.

In the event of a month-long gap in the observation data, a default model sampling strategy is required. In this case, model samples are chosen from the same hours that an observation sample would have been taken (daily selected hours for quasi-continuous sites and on five days during the month for flask sampling sites). The default hours for flask sampling are based on the distribution of hours when samples were taken during 2000-2004 at the respective sampling sites. Monthly means computed from these selected hours are used to fill in the gaps in the co-sampled transport matrix $\mathbf{H}$.

2.3. Results

2.3.1. Global Results

Figure 2-3 shows the mean annual sources and analytically-derived uncertainties for the analysis period (2001-2003) for the three networks tested (triangles for the base network, diamonds for the enhanced network, and squares for the continental extension), aggregated to the 22 TransCom continent and ocean basin regions. Results for the continental extension network are also listed in **Table 2-1** and **Table 2-2**. All results are adjusted fluxes including the fixed

biomass burning emissions, but not the fossil fuel emissions. This convention allows comparisons to previous work in the literature. Error bars in these figures are ± the mean annual uncertainty from the posterior covariance matrix, taking into account off-diagonal values wherever the results are aggregated from the region-month level.

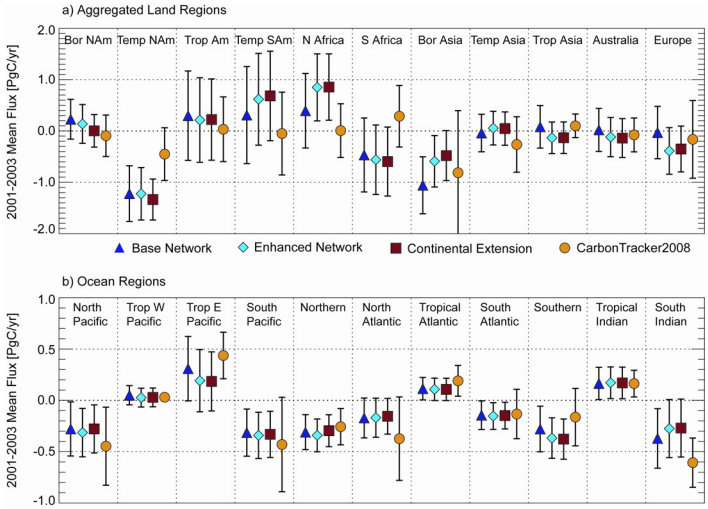

Figure **2-3**: 2001-2003 mean annual estimated flux and uncertainty in Pg C yr^{-1} for a) aggregated land and (b) ocean regions for 3 networks: 54-site base network (triangles), 86-site enhanced network (diamonds) and 91-site continental extension network (squares). CarbonTracker 2008 results for the same time period are shown for reference (circles). Fossil emissions are not included.

The CarbonTracker 2008 (http: //carbontracker.noaa.gov) results for the same time period are shown as a benchmark reference (circles in **Figure 2.3**). For all three networks in our experiment the increase in the global atmospheric burden of carbon (from the net of the estimated global sink and the assumed biomass burning and fossil fuel emissions) is consistent within 0.1 Pg C yr^{-1} with the ~+2.1 ppm average annual increase in global surface carbon dioxide for this period. The proportion of the global sink attributed to land regions across the three networks (54% for the base network, 44% for the enhanced network, and 47% for the continental extension), shows, at the global level, some sensitivity of the solution to the composition of the

Table **2-1**: Spatial resolution of the inversion with prior specifications and 2001-2003 mean annual posterior flux solution for the continental extension network. All units are Pg C yr^{-1}. Net Post. Flux includes biomass burning emissions, but not fossil emissions.

	Annual Mean Prior Specifications				Annual Mean Posterior Results		
	Prior Flux	Prior Error	Fossil Emiss.	Burning Emiss.	Post. Flux	Net Post. Flux	Post. Error
Land Regions							
Boreal North America							
Western Boreal	0.00	1.09	0.01	0.01	-0.16	-0.15	0.36
Northern Boreal	0.00	0.22	0.00	0.00	0.02	0.02	0.18
Eastern Boreal	0.00	0.62	0.00	0.01	0.12	0.13	0.22
Temperate North America							
Pacific Northwest	0.00	0.47	0.08	0.01	-0.43	-0.42	0.35
Central Plains	0.00	0.55	0.14	0.00	0.09	0.09	0.20
North Central	0.00	0.30	0.13	0.00	-0.10	-0.10	0.13
Northeast	0.00	0.84	0.78	0.00	-0.72	-0.72	0.24
Southwest	0.00	0.16	0.30	0.00	-0.10	-0.09	0.08
Southeast	0.00	0.37	0.38	0.00	-0.08	-0.07	0.25
Subtropical	0.00	0.21	0.07	0.00	-0.03	-0.02	0.19
Tropical America							
Central America	0.00	0.43	0.11	0.05	0.01	0.06	0.35
Northern Amazon	0.00	0.75	0.03	0.04	-0.07	-0.04	0.49
Southern Amazon	0.00	0.90	0.00	0.11	0.09	0.19	0.75
Temperate South America							
Northern	0.00	1.53	0.12	0.12	0.56	0.68	0.89
Southern	0.00	0.27	0.01	0.00	0.00	0.00	0.25
Northern Africa							
Mediterranean Coast	0.00	0.14	0.05	0.00	0.11	0.11	0.11
Northern Arid	0.00	0.14	0.04	0.00	-0.03	-0.03	0.13
Northern Dry Savanna	0.00	0.67	0.01	0.12	-0.04	0.08	0.58
Northern Mesic Savanna	0.00	1.10	0.02	0.48	0.22	0.70	0.73
Horn of Africa	0.00	0.17	0.00	0.01	0.00	0.00	0.17
Southern Africa							
Western Forest	0.00	0.43	0.02	0.06	-0.07	-0.01	0.42
Southern Mesic Savanna	0.00	1.39	0.01	0.46	-0.93	-0.47	0.63
Southern Dry Savanna	0.00	0.34	0.10	0.06	-0.16	-0.09	0.27
Southern Arid	0.00	0.14	0.00	0.00	-0.04	-0.04	0.14
Madagascar	0.00	0.14	0.00	0.02	-0.01	0.01	0.14
Boreal Asia	0.00	3.42	0.14	0.21	-0.69	-0.48	0.49
Temperate Asia	0.00	0.95	2.16	0.04	0.01	0.05	0.33
Tropical Asia	0.00	0.47	0.53	0.24	-0.37	-0.13	0.31
Australia	0.00	0.51	0.10	0.16	-0.30	-0.14	0.38
Europe							
British Isles	0.00	0.14	0.18	0.00	-0.01	-0.01	0.12
Scandinavia	0.00	0.36	0.08	0.00	0.01	0.01	0.14
North Central	0.00	1.69	0.49	0.03	-0.83	-0.80	0.55
Western	0.00	0.28	0.31	0.00	0.03	0.03	0.15
Interior Central	0.00	0.39	0.41	0.00	-0.08	-0.08	0.17
Eastern	0.00	0.71	0.32	0.01	0.45	0.46	0.34
Iberia	0.00	0.16	0.09	0.00	0.04	0.04	0.14

Table 2-1 (continued)

	Annual Mean Prior Specifications				Annual Mean Posterior Results		
	Prior Flux	Prior Error	Fossil Emiss.	Burning Emiss.	Post. Flux	Net Post. Flux	Post. Error
Ocean Regions							
North Pacific	-0.49	0.53	0.02	0.00	-0.28	-0.28	0.23
Tropical West Pacific	0.11	0.10	0.00	0.00	0.03	0.03	0.09
Tropical East Pacific	0.51	0.44	0.00	0.00	0.18	0.18	0.29
South Pacific	-0.26	0.29	0.00	0.00	-0.33	-0.33	0.23
Northern	-0.29	0.25	0.00	0.00	-0.30	-0.30	0.16
North Atlantic	-0.28	0.32	0.02	0.00	-0.16	-0.15	0.17
Tropical Atlantic	0.13	0.11	0.00	0.00	0.11	0.11	0.11
South Atlantic	-0.13	0.14	0.00	0.00	-0.15	-0.15	0.13
Southern	-0.65	0.58	0.00	0.00	-0.38	-0.38	0.20
Tropical Indian	0.19	0.18	0.00	0.00	0.17	0.17	0.15
South Indian	-0.48	0.44	0.00	0.00	-0.27	-0.27	0.29

network. Although the tropical and southern regions are not individually well-constrained, for all three networks the tropical land/ocean source and the southern land/ocean sink net to a 0.2 Pg C yr^{-1} source. The tropical and southern regions as a latitudinal band are constrained by these networks, but the flux distribution among these regions is not well-defined. The ocean region fluxes are mainly consistent within uncertainties across the three networks.

In the northern regions, the three networks attribute 71-74% of the northern sink to the land regions, with the fluxes for the ocean regions (North Pacific, North Atlantic and Northern Oceans) consistent across the three solutions. However, the distribution of the land fluxes among continents depends very much on the network. For the base network, the land sink is split between Asia (Boreal and Temperate Asia regions accounting for 51%) and North America (Boreal and Temperate North America regions accounting for 47%) leaving Europe essentially neutral. Introducing east Asian observation sites, the western Pacific high altitude aircraft observations and mountain-top observations in Europe in the enhanced network shifts some of the Asian flux to Europe for an Asia:North America:Europe balance of 27%:54%:19%. This flux

redistribution is accompanied by reductions in uncertainty for Europe and Asian regions (for example, Tropical Asia region annual uncertainty changes from 0.41 to 0.31 Pg C yr^{-1} and Europe from 0.51 to 0.39 Pg C yr^{-1}). The continental extension network with the four additional continental North American sites shifts more of the Asian flux to North America (20%:63%:17%), with a reduction in uncertainty in North America (from 0.46 to 0.38 Pg C yr^{-1}), but not in Asia or Europe. The absence of observation sites in Boreal Asia and North Central Europe makes these regions under-constrained.

Two network-related shifts are seen in ocean fluxes between the base network and the enhanced network: the first is a transfer of source from the Tropical East Pacific and Tropical America to Temperate South America; the second is a shift of the distribution of the sink between the Southern Ocean and the South Indian Ocean with only modest reductions in uncertainty. Southern Hemisphere observation site additions in the enhanced network include Jubany Bay in Antarctica, Cape Point in southern Africa, and the southern branch of the high altitude western Pacific aircraft flights. The flux tower in the Amazon in the continental extension network makes little difference in either flux or uncertainty in the aggregated South American regions. The modeled atmospheric concentration of carbon dioxide at this site fits the data poorly before the inversion and the post-inversion predicted concentration here is the worst fit of all of the observation sites used. This issue of correctly modeling the terrestrial carbon flux in the Amazon continues to be addressed (Baker et al., 2008). The poor model fit to the site observation and the lack of any other observations on the South American continent likely account for the lack of improvement in the certainty of estimated fluxes here.

The mean annual uncertainty reductions ($1 - \sigma_{post} / \sigma_{prior}$, expressed as a percent, where σ_{post} is the 2001-2003 mean annual posterior uncertainty and σ_{prior} is the annualized prior uncertainty) for a subset of the regions, those in North and South America, are shown in **Figure 2-4** for the base network (triangles), enhanced network (diamonds), and continental extension

network (squares). The annualized prior uncertainty and mean annual posterior uncertainty for each region are also listed in **Table 2-1**. **Figure 2-4** demonstrates that uncertainty reduction is

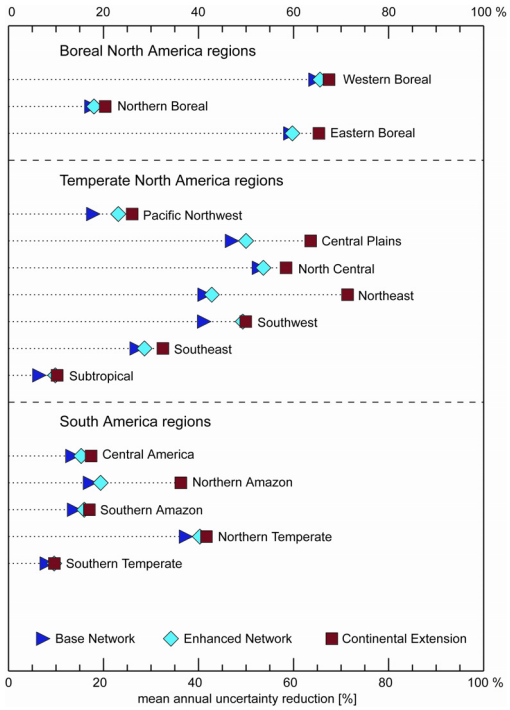

Figure **2-4**: Mean annual uncertainty reduction for 2001-2003 (in per cent) for North and South American regions for three networks: base network (triangles), enhanced network (diamonds), and continental extension network (squares).

achieved in the regions local to the added observations, a finding consistent with results from grid-scale inversions (e.g., Peylin et al., 2005; Mueller et al., 2008; Gourdji et al., 2008; Lauvaux et al., 2008). In regions where there are no observation sites in any of these three networks (for example, the Subtropical region within Temperate North America and the Southern region within Temperate South America), there is no reduction in uncertainty. Introducing one or two sites in a previously unrepresented region can result in a significant reduction in uncertainty. The Northeast region within Temperate North America shows little improvement in the enhanced network compared to the base network (43% vs. 41%); adding two flux tower observation sites in the

continental extension network results in an uncertainty reduction of 71% from the prior uncertainty with a mean annual posterior uncertainty of 0.24 Pg C yr^{-1} compared to 0.48 Pg C yr^{-1} for the enhanced network. The South American flux tower is on the border between the Northern and Southern Amazon regions within Tropical America. The percentage reduction in uncertainty in the Northern Amazon compared to the prior doubles in the continental extension network compared to the base and enhanced networks (17% for the base network, 19% for the enhanced network, 36% for the continental extension). The Southern Amazon region shows no improvement at all. Parazoo et al. (2008), using the same transport model and analyzed meteorological fields, found that atmospheric mixing in this region is dominated by vertical convection rather than the horizontal transport by synoptic weather systems prevalent in the mid-latitude continents. It is not surprising, therefore, that one observation site in South America constrains only the immediate region.

Regions already partially constrained by local observations show some improvement, but less than for the first observation site. The Southwest region within Temperate North America is the only region shown in this figure with observations introduced in the enhanced network. Adding the aircraft vertical profile observations at Carr, Colorado, results in a modest improvement from 41% to 49% from the prior specification. A flux tower observation site is added in the continental extension network to each of two regions with an existing observation. In Eastern Boreal North America the percentage reduction of the uncertainty improves from 60 to 65%, and for the Central Plains the change is from 50 to 64%.

The posterior covariance matrix for 2002 annual fluxes for the aggregated TransCom regions and the continental extension network is presented in **Figure 2-5**. Variances are reported on the diagonal. Off-diagonal values are annual covariances indicating the extent to which the inversion determines the adjustments to the regional annual fluxes as independent on an annual basis. A fully independent solution would have no off-diagonal entries in the posterior

covariance. Shaded off-diagonal values, in variance units of $(\text{Pg C yr}^{-1})^2$, impact the annual uncertainty of the aggregated regions by 0.1 Pg C or more. The ocean posterior estimates

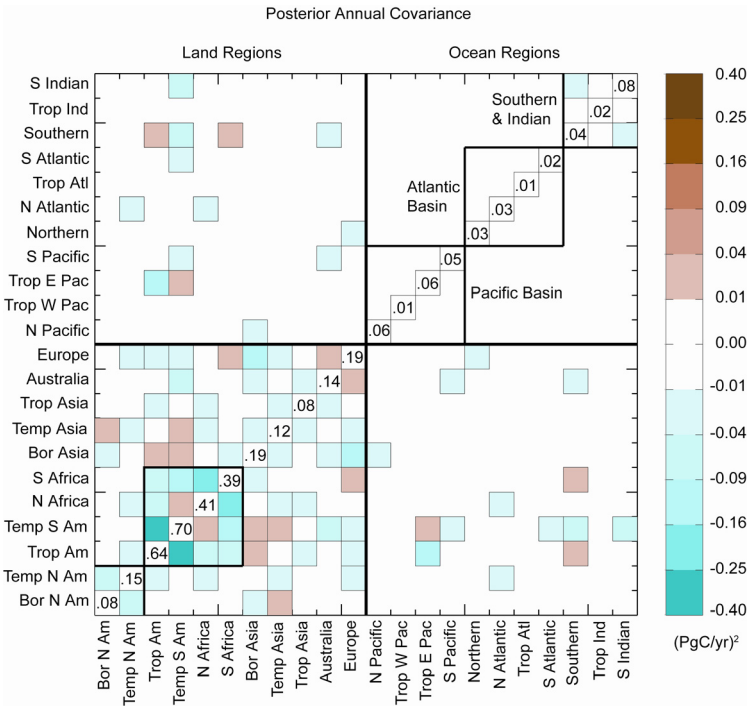

Figure **2-5**: Posterior annual covariance matrix for 2002 for the continental extension network solution for the 22 aggregated continental and ocean basin regions. Variances in $(\text{Pg C yr}^{-1})^2$ are shown on the diagonal. Shaded boxes contribute $\geq$ 0.1 Pg C yr^{-1} to the aggregated region uncertainty.

can be considered independent by this measure, with the minor exception of the relationship between the Southern Ocean and the South Indian Ocean. The largest covariances are between Tropical America and Temperate South America (-0.38 in units of variance), North Africa and South Africa (-0.21), Temperate South America and South Africa (-0.13), and Europe and Boreal Asia (-0.10). Indeed, the negative covariances among the four South American and African aggregated regions, bounded by heavy black lines in **Figure 2-5**, indicate that the estimates for

the two continents should not be treated as independent. The total flux may be constrained, but the partitioning is not certain. Introducing the observation site in South America, where there were none in the enhanced network, is not enough to improve the overall continental constraint. In North America, however, the four added observations reduce the 2002 annual variance for Boreal North America from 0.12 to 0.09 (Pg C yr^{-1})^2, for Temperate North America from 0.26 to 0.15 (Pg C yr^{-1})^2 and the covariance between them from -0.09 to -0.05 (Pg C yr^{-1})^2 (bounded by heavy black lines in the lower left corner of **Figure 2-5**).

Similar improvements are seen in the non-aggregated, 47-region 2002 annual covariance matrix for the continental extension network (not shown). The covariances among the regions within North America are reduced; the covariance between Eastern Boreal and Northeast Temperate regions is reduced from -0.05 to -0.02 (Pg C yr^{-1})^2, and the variances for Central Plains and Northeast are improved (0.09 to 0.04 (Pg C yr^{-1})^2 for the Central Plains and 0.22 to 0.05 (Pg C yr^{-1})^2 for the Northeast). Covariances between the North American regions and non-North American regions are reduced except for a negative covariance between Western Boreal North America and Boreal Asia. Results are more modest in South America, where the only appreciable improvement is the reduction in the covariance between Northern and Southern Amazon regions from -0.11 to -0.07 (Pg C yr^{-1})^2. Covariances within South America and between South American and African regions are reduced but not eliminated. We cannot justify the inversion solution for the sub-regions within South America as independent with the networks tested.

2.3.2. North American Results

Figure 2-6 shows the 2001-2003 mean annual North American fluxes and uncertainties along with the bordering North Atlantic and North Pacific ocean regions as well as Europe and Boreal Asia. Here it is evident that the uncertainty reduction in the regions with the added sites in

the continental extension network (Eastern Boreal North America, Central Plains and Northeast in Temperate North America) accounts for all of the uncertainty reduction in North America seen in **Figure 2-3**. Surprisingly, the source/sink values change little between the enhanced and continental extension networks, unlike the change in balance between Boreal Asia and Europe when adding the European mountain top and additional sites in the eastern Asia. Uncertainty in the regions upstream of the added observations in the northern tier (Western Boreal and Pacific Northwest) is not much improved. This is also evident in the uncertainty reduction percent shown in **Figure 2-4**.

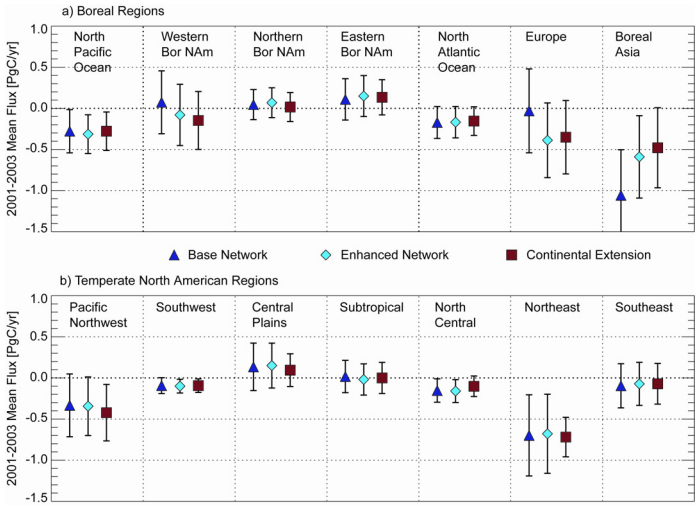

Figure **2-6**: Mean annual fluxes and uncertainties for 2001-2003 for the regions within a) Boreal North America and b) Temperate North America for three networks: base network (triangles), enhanced network (diamonds) and continental extension network (squares). Also shown in a) are other northern latitude aggregated regions. All units are Pg C yr^{-1}.

The monthly flux solutions for the North American regions for the three networks are shown in **Figure 2-7**. The base network results are shown in blue, the enhanced network in cyan and the continental extension in red. Dotted lines indicate the extent of the constraint provided by the prior uncertainty. Shading represents the posterior uncertainty (from the diagonal of the posterior covariance matrix) of the solution for the continental extension network. The top row of

the figure shows the regions within Boreal North America followed by the aggregated region results. The second and third rows show the regions within Temperate North America followed by its aggregated result. The greatest effect of the continental extension network can be seen in

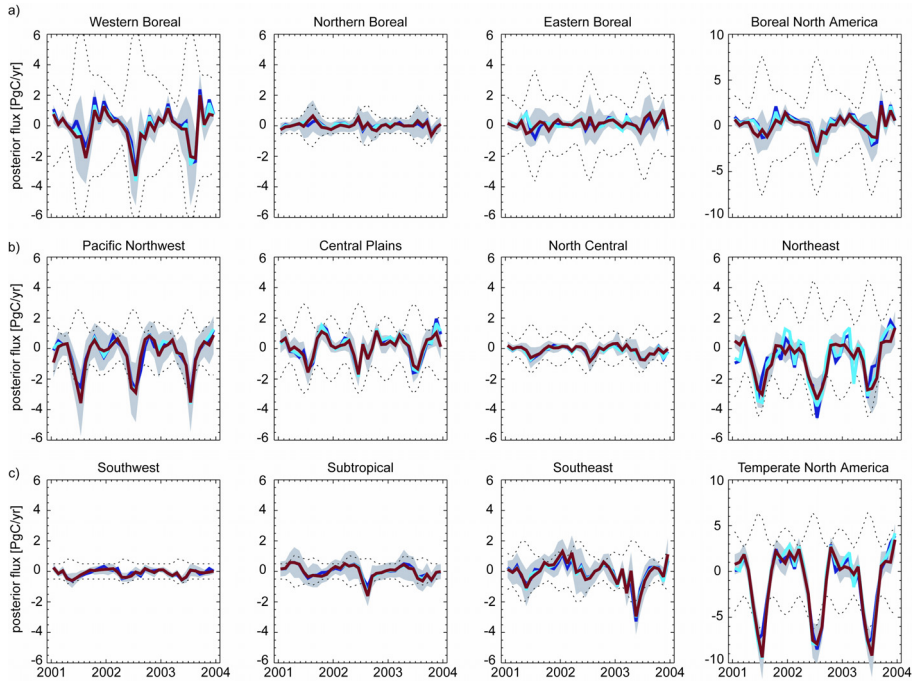

Figure **2-7**: Monthly posterior fluxes (Pg C yr^{-1}) for a) regions within Boreal North America and b) and c) Temperate North America. The aggregated regions (not different scales) are at the end of rows a) and c). Base network results are in blue, enhanced network in cyan, and continental extension network in red. Shading is posterior uncertainty of continental extension network. Dotted lines define the prior uncertainty.

the Northeast region of Temperate North America where two flux towers with well-calibrated CO_2 measurements are located. The posterior source/sink solution is smoothed, of smaller amplitude relative to the prior, and has a reduced posterior uncertainty. The other new towers in this network are in the Central Plains region of Temperate North America and the western edge of the Eastern Boreal region; each of these regions also contains one other observation site, the WKT tall tower in the Central Plains region and the Fraserdale site in the Eastern Boreal region. The areas with the least uncertainty reduction (Northern Boreal, Pacific Northwest, Southeast,

and Subtropical) either have no observation sites or one site at the edge of the region (Alert in Northern Boreal and Key Biscayne on the southeastern Florida coast). The posterior uncertainty estimated for Northeast and Central Plains is small relative to the prior uncertainty, showing again that the impact of the added observation sites is primarily local. The overall effect of the new towers on the aggregated regions is to reduce the uncertainty and to smooth some of the noise in the monthly solution.

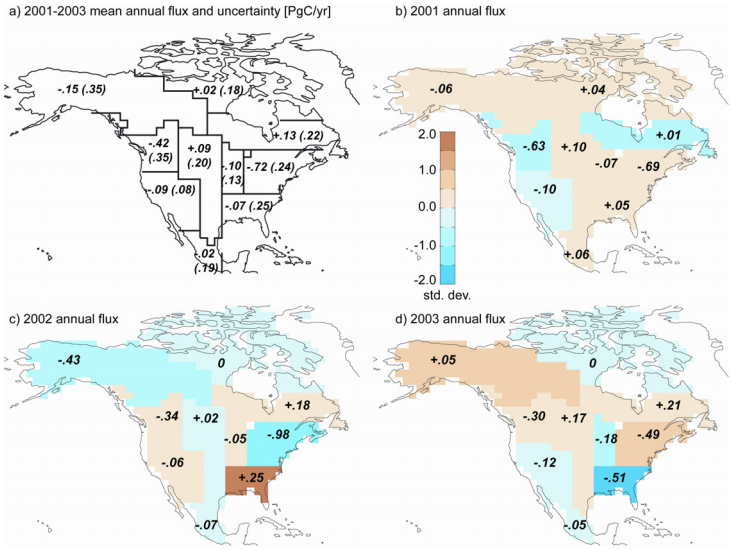

Figure **2-8**: a) 2001-2003 mean annual flux (uncertainty) in Pg C yr^{-1} for North American regions. b), c), d) show annual results for 2001, 2003 and 2003. Shading on the annual maps indicates the number of standard deviations from the mean annual flux for the year as an indication of interannual variability.

Figure 2-8 shows the interannual variability of the source/sink solutions for the continental extension network for the North American regions. 2001-2003 mean annual flux and mean annual uncertainty (in parentheses) in Pg C yr^{-1} are shown by region in panel a. The other panels show the results for each of the years 2001 through 2003. Shading indicates the fractional standard deviation of the annual flux from the three-year mean. The southeast region has the largest interannual variability with a strong source deviation with severe drought in summer of

2002 (Waple et al., 2003), but a strong sink anomaly in with the wet summer of 2003 (Levinson and Waple, 2004). The variability in the northeast region exhibits the reverse behavior. The 47-region covariance matrix shows a negative annual covariance of -0.01 $(Pg\ C\ yr^{-1})^2$ between these two regions. There are no observation sites in the southeast region, so it is possible that this dipole behavior is an artifact of flux balancing in the inversion. All other anomalies are within 1 standard deviation of the three-year mean.

2.4. Discussion

2.4.1. Impact of Added Measurement Sites

The added continental measurement sites contribute to local uncertainty reduction without undue disruption to the global flux solution. The mid-day measurements appear to be compatible with the transport model and the spatial resolution of the inversion solution in this case. As these hours are most likely to represent the desired well-mixed background state, they are likely to represent an area much broader than the immediate vicinity of the observation site. This avoids the issue of nocturnal boundary layer representation in the transport models and driver meteorological data. Representation error is still likely; choice of site is critical. The competing roles of multiple sites in the same solution region (Howland Forest and Harvard Forest in the Northeast region of Temperate North America) warrant further research. There are a number of additional North American flux towers whose carbon dioxide measurements are now well-calibrated. They present an opportunity for improving the uncertainty of the solution with the addition of these continental sites. The use of these stations is, however, dependent on the continuity of the measurement time series, and the degree to which the site is representative of the area for flux estimation. The inversion method used here requires continuous observation time

series and an unchanging network. Other methods may be more amenable to the addition and deletion of observation sites for an inversion solution for long periods of time.

2.4.2. Comparison to Published Results

The results of this inversion are placed in the context of another contemporary inversion for the same time period in **Table 2-2** and **Figure 2-9** (see also the CarbonTracker 2008 results shown for reference in **Figure 2-3**). Also shown is the NASA GSFC PCTM submission to the TransCom Interannual Variability (IAV) experiment, extended forward in time to participate in the North American Carbon Program Interim Synthesis activity (A. Jacobson, personal communication). The TransCom IAV-PCTM inversion shares the same tracer transport model

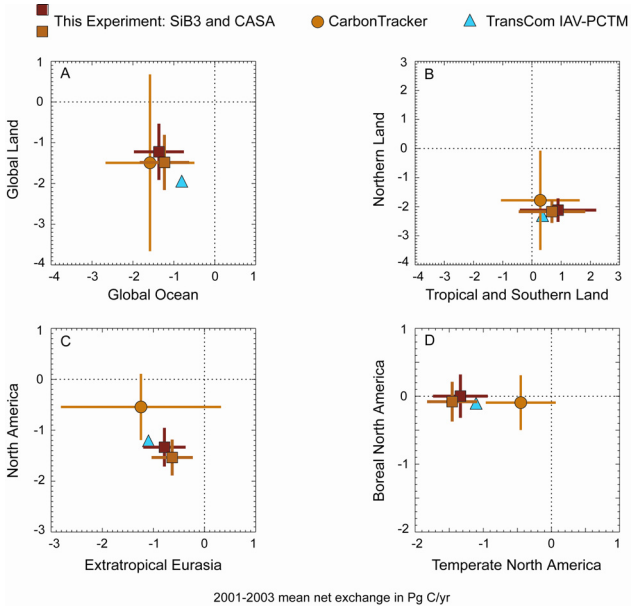

Figure **2-9**: Comparison of partitioning of annual mean fluxes for 2001-2003. Results are from this study (squares) using both SiB3 and an alternative background terrestrial flux (CASA climatology), CarbonTracker 2008 (circles) and the NASA PCTM inversion from the TransCom Interannual Variability experiment (triangles)

with our experiment, but uses a repeated year of meteorological driver data, a smoothed monthly data product for observation data, a different observation network, and no explicit handling of biomass burning emissions. Our results (squares in **Figure 2-9**) include the continental extension network solution discussed in this paper and an inversion solution accomplished by substituting the monthly mean CASA flux instead of the hourly, interannually variable SiB3 flux as the background terrestrial flux.

Table **2-2**: Comparisons of 2001-2003 mean annual fluxes (Pg C yr^{-1}) from this study using SiB3 and CASA background terrestrial fluxes, the NASA PCTM inversion solution from the TransCom Interannual Variability experiment and CarbonTracker2008. Uncertainty information is not available for the NASA PCTM solution. Fossil fuel emissions are not included.

	This Study SiB3 bio	This Study CASA bio	TransCom IAV PCTM	Carbon Tracker
Aggregated Land Regions				
Boreal North America	0.00 ± 0.32	-0.08 ± 0.29	-0.12	-0.07 ± 0.40
Temperate North America	-1.34 ± 0.40	-1.46 ± 0.37	-1.11	-0.47 ± 0.51
Tropical America	0.22 ± 0.79	-0.59 ± 0.61	0.64	0.05 ± 0.63
Temperate South America	0.68 ± 0.87	0.61 ± 0.82	-1.32	-0.07 ± 0.81
Northern Africa	0.86 ± 0.65	0.99 ± 0.62	0.81	0.01 ± 0.52
Southern Africa	-0.60 ± 0.67	-0.24 ± 0.64	0.07	0.27 ± 0.60
Boreal Asia	-0.48 ± 0.49	-0.72 ± 0.44	0.48	-0.69 ± 1.21
Temperate Asia	-0.05 ± 0.33	0.35 ± 0.34	-0.69	-0.25 ± 0.54
Tropical Asia	-0.13 ± 0.31	-0.15 ± 0.30	0.33	0.10 ± 0.23
Australia	-0.14 ± 0.38	0.06 ± 0.36	-0.19	-0.07 ± 0.33
Europe	-0.35 ± 0.45	-0.27 ± 0.39	-0.92	-0.16 ± 0.76
Ocean Regions				
North Pacific Ocean	-0.28 ± 0.23	-0.38 ± 0.23	-0.72	-0.43 ± 0.38
Tropical West Pacific Ocean	0.03 ± 0.09	0.05 ± 0.09	-0.07	0.03 ± 0.02
Tropical East Pacific Ocean	0.18 ± 0.29	0.33 ± 0.28	0.33	0.44 ± 0.23
South Pacific Ocean	-0.33 ± 0.23	-0.29 ± 0.23	-0.01	-0.41 ± 0.46
Northern Ocean	-0.30 ± 0.16	-0.21 ± 0.15	-0.09	-0.26 ± 0.18
North Atlantic Ocean	-0.15 ± 0.17	-0.14 ± 0.17	-0.20	-0.40 ± 0.41
Tropical Atlantic Ocean	0.11 ± 0.11	0.11 ± 0.11	0.19	0.19 ± 0.15
South Atlantic Ocean	-0.15 ± 0.13	-0.20 ± 0.13	-0.12	-0.12 ± 0.24
Southern Ocean	-0.38 ± 0.20	-0.37 ± 0.20	-0.11	-0.19 ± 0.28
Tropical Indian Ocean	0.17 ± 0.15	0.14 ± 0.15	0.09	0.16 ± 0.13
South Indian Ocean	-0.27 ± 0.28	-0.26 ± 0.28	-0.08	-0.60 ± 0.24

Figure 2-9 shows that our experiment and CarbonTracker 2008 (results obtained from http: //carbontracker.noaa.gov) both partition the global sink approximately equally between land and ocean (panel A), and that the global land sink in the northern land regions is partly offset by a source in the aggregated tropical and southern land regions (panel B). In contrast to CarbonTracker, however, our results place more of the northern land sink (panel C) in North America than in extratropical Eurasia (Europe, Boreal Asia and Temperate Asia), and more of the North American land sink (panel D) in Temperate than in Boreal North America. Only the CarbonTracker results in panel C are in line with an independent estimate of -0.50 Pg C/yr (not including fossil fuel emissions) (CCSP, 2007) for North America. The partitioning behavior at the aggregation level shown in **Figure 2-9** for our experiment is similar to that of the TransCom IAV-PCTM results. In this experiment the influence of the transport model appears to dominate over the network and the terrestrial background flux used. It is possible that using mid-day samples from the continental observation sites introduces a negative bias in the observation data during the growing season that could account for the increase in the North American sink in our experiment compared to the TransCom IAV-PCTM result. This could be tested by comparison with an inversion using vertical column observations of carbon dioxide from satellite measurements.

The differences between the SiB and CASA versions of our experiment are small relative to the uncertainties for these continental aggregated areas. Other studies have suggested that the transport model used and the uneven and sparse spatial distribution of observation sites are equal contributors to the differences in results (Gurney et al., 2002; Gurney et al., 2008; Roedenbeck et al., 2003a; Roedenbeck et al., 2003b). Here we can compare inversions using different background fluxes and the same network (the SiB3 and CASA versions of our experiment), inversions using the same transport model and inversion method, but different networks (our experiment and the TransCom IAV-PCTM solution), and inversions using different methods,

different transport models, and different networks of observations (our experiment and CarbonTracker). Since the solution at the global level should be determined by hemispheric-scale gradients and global growth rates, the results should agree at the large scale as seen in **Figure 2-9** panels A and B. At the finer scale, the differences due to observation network distribution and density are more evident as can be see in **Table 2-1** and **Figure 2-3**. **Figure 2-9** panel D, in particular, shows that similar reductions in uncertainty can be achieved by very different methods for well-constrained regions such as North America. Networks and transport matters more than background fluxes in this experiment.

2.4.3. Adjustments of the Background Fluxes

Our inversion results are generally consistent in zonal distribution with the ocean air-sea flux estimates of Takahashi et al. (2002), which we use as our background ocean flux. Our flux solution reduces the net global ocean sink from 1.6 Pg C to 1.5 Pg C annually, with smaller sinks in temperate waters and smaller sources in the tropics. Our experiment reduces the Southern Ocean sink specified by Takahashi et al. (2002), but not by as much as CarbonTracker (see Table 2-2) or the update to these air-sea fluxes (Takahashi et al., 2009) for the climatologic year 2000, which shows little or no net annual flux in the Southern Ocean due to offsetting effects of biological drawdown and upwelling of carbon-rich waters.

We also find a reduction in amplitude of the seasonal cycle in the Northern Hemisphere boreal and northern temperate land region fluxes compared to SiB3, where our results are reasonably well-constrained by observations. This can be seen in the prior uncertainty ranges (derived from the SiB3 fluxes) in **Figure 2-7**. The CASA terrestrial background flux has smaller annual amplitude in these same regions, but as can be seen in **Figure 2-9**, yields an inversion solution very similar to that using SiB3 in regions that are well-constrained. The fit of predicted

observations to the data used in the inversion is very good for remote observation sites, and generally good overall. This is true for both the SiB3 and CASA versions of the inversion.

2.5. Conclusions

We have shown the value of additional continental observations in a standard Bayesian global atmospheric inversion. Specific findings include the following:

1. Using a judicious combination of region size and location of additional continental observation sites, the posterior flux uncertainty and region-to-region covariance are significantly reduced.

2. Adding observation sites in unconstrained regions will further reduce the uncertainties. Flux towers are typically sites in representative ecosystems and have existing measurement infrastructure. With some additional instrumentation and a careful calibration of carbon dioxide concentrations to global standards, these sites can be used in the global measurement network used for atmospheric inversions. With the extent of the current global flux network, there is a major opportunity for improved coverage in the observation networks used for global and regional atmospheric inversions. The use of time series from these flux towers is contingent on operational funding to continue long time series of carbon dioxide observations. Operations have been discontinued at two of the five flux towers used in this experiment.

3. Adding observation sites in regions already constrained may push the boundaries of this methodology and deserves additional research on balancing region size with representative ecosystems sampled by the observing sites.

4. The transport model and analyzed meteorological fields appear to have more impact on our flux estimates than the added observation sites or the background terrestrial flux. For the well-constrained aggregated regions, the TransCom-IAV PCTM results are more like our findings

than CarbonTracker. Transport uncertainty is an important problem that this study does not address.

5. We confirm the findings of other researchers (e.g., Baker et al., 2006; Jacobson et al., 2007a; Jacobson et al., 2007b) that the tropical and southern land regions are under-constrained to the extent that the inversion source/sink estimates for these regions cannot be considered independently. The observation networks tested in this experiment do not improve this situation. An increase in observation density, for example from satellite observations, is required to change this situation.

Chapter 3

Sensitivity to Methods Choices

3.1. Introduction

We have chosen to treat this problem as linear using a least squares, analytic solution method; the carbon dioxide mixing ratio at each observation point and month is treated as a linear combination of responses to surface-atmosphere fluxes from the regions in the months defined in our solution space. This least squares method requires that the prior uncertainties be modeled using Gaussian distributions (Tarantola, 2005), both the data uncertainties (combined measurement and transport errors in our case) and prior model uncertainties (the prior flux uncertainties). The advantage of oversimplifying the errors in this way is a relatively simple solution method; the main disadvantage is a lack of robustness. The solution may be particularly sensitive to outlier data points (Tarantola, 2005).

We have also attributed all changes to carbon dioxide in the atmosphere to surface fluxes. There is an inherent bias in the assignment of all atmospheric carbon dioxide sources to surface grid boxes and months (Suntharalingam et al., 2005). Precursor trace gas (carbon monoxide, methane, and non-methane volatile organic compounds) are emitted from the surface, and the carbon dioxide is an oxidation product in the atmosphere downwind in space and later in time from the region and time of surface emission. Suntharalingam et al. (2005) compared the results of three annual mean inversions to the results obtained using a transport model including the appropriate atmospheric chemistry and found that such an assumption causes the northern land sink to be overestimated by ~0.25 Pg C yr^{-1} and tropical efflux to be overestimated by a smaller amount. We have not addressed this bias in our experiment.

Methodological advances introduced in this experiment include using actual data for observations rather than a smoothed data product and introducing annual variability in the transport fields used to generate the response functions. These are departures from the TransCom inversion method on which this experiment is based (e.g., Baker et al., 2006). Important assumptions related to the observation data include independence of each observation datum and Gaussian distribution of the data uncertainties, which for this method include transport error as well as measurement error. This independence cannot be true, as the observation data set includes co-located continuous and discrete sampling sites, and the monthly mean observations exhibit seasonal autocorrelation. Tarantola (2005) cautions us to examine the inversion results to see how far we have strayed from valid use of this least squares solution method. The focus of this chapter is examination of the inversion solution for obvious effects of non-compliance.

Section 3.2 documents positive controls on the inversion method. In Section 3.3 we compare inversion results using different choices of the terrestrial and fossil fuel background fluxes to test for over-reliance on the background fluxes. In Section 3.4 we show what happens when the prior flux uncertainties are modified, removing restrictions that limit the departure from the background fluxes. In Section 3.5 we examine a possible enhanced sensitivity to outliers, by showing an inversion result for a network that includes one station not used in the continental extension network. We also try to determine if the two sites in the Northeast region of Temperate North America provide redundant information to the inversion. In Section 3.6 we explore the suggestion in Tarantola (2005) of consistency checking the solution for the importance of the assumption about Gaussian errors. The analysis of the posterior covariance in Chapter 2 is another of these checks; this will be revisited in Chapter 5. We partially address the problem of highly correlated observations using an inversion variation with no co-located observation sites. In Section 3.7 we will examine the data residuals and introduce 2 more inversion variations with additional model-data mismatch error added to the observation data uncertainties. Section 3.8

concludes the chapter with an overall assessment of the use of this method. The test of sensitivity to other network design choices besides the co-located sites will be addressed in Chapter 4.

The inversion variations discussed in this chapter are documented in **Table 3-1**. Observation Site-Network cross references for these inversions are in the Appendix in **Table A-9**.

Table **3-1**: Inversion variations discussed in Chapter 3.

Inversion	Network	# of Sites	Description
Control Inversion	CE	91	Continental Extension Network inversion described in Chapter 2; model output co-sampled; uses SiB3 terrestrial and seasonal fossil background fluxes
Positive Control Tests			
Posterior flux as prior	CE	91	Posterior flux estimates from the CE network inversion are substituted for the prior fluxes
Posterior flux and uncertainty as prior	CE	91	Posterior flux and uncertainty estimates from the CE network inversion are substituted for the prior fluxes and uncertainties
Perfect data & perturbed perfect data	CE	91	Model samples from the forward integrations of the background fluxes (and perturbed model samples) are substituted for the observation data
Sensitivity to Choice of Background Fluxes			
SiB, annual fossil	CE	91	SiB3 terrestrial and annual fossil background fluxes
CASA, seasonal fossil	CE	91	CASA terrestrial and seasonal fossil background fluxes
CASA, annual fossil	CE	91	CASA terrestrial and annual fossil background fluxes
Sensitivity to Prior Flux Uncertainty			
TransCom prior	CE	91	Control inversion with prior flux uncertainty from the TransCom IAV experiment (Gurney et al., 2008)
Uncertainty x n, where n=0.5, 1, 2, 5, 10, 50	CE	91	Control inversion with monthly prior uncertainty magnitudes changed universally by factors of 0.5, 2, 5, 10, 50
Sensitivity to a Single Observation Site			
CE with ESP	N5	92	CE network plus Estevan Point
Enhanced Network Inversion	E	86	Enhanced network inversion from Chapter 2 (CE network without the 5 flux towers)
CE without HFM	N6	90	CE network without Harvard Forest
CE without HOW	N7	90	CE network without Howland Forest
Sensitivity to Co-Located Sites			
No co-located sites	N8	78	CE network with co-located surface sites removed
Model-Data Mismatch Error Tests			
Mismatch test 1	N8	78	Data uncertainty increased (level 1)
Mismatch test 2	N8	78	Data uncertainty increased (level 2)

3.2. Positive Controls on the Inversion Method

In any complex modeling experiment, it is always wise to execute controlled tests where the results can be compared to expectations. We conduct tests both from the perspective of perfect fluxes and of perfect data. For the perfect flux experiments we replace the prior fluxes in an inversion using the 91-site CE (continental extension) network with the flux adjustments calculated in the CE network inversion. In the original inversion the posterior flux solution is a sum of the background fluxes, the prior fluxes (specified as zero), and the calculated adjustments. For the first perfect flux test, we substitute the calculated adjustments for the prior fluxes, and expect that the inversion calculation will result in zero additional adjustments given the same observation network. In the second test, the prior specifications include both the calculated adjustments and the posterior flux uncertainties. This inversion should also result in zero additional adjustments, but with a narrower posterior uncertainty.

Figure 3-1 shows the mean annual results for 2001-2003 of the perfect flux tests for the aggregated North American region (panel A), and three of the North American sub-regions: Eastern Boreal, Pacific Northwest, and Northeast Temperate. The inversion method dictates that if the *a priori* specifications have Gaussian distributions, then the posterior solution can also be described as a Gaussian. For example, prior flux specifications in **Figure 3-1** (black dashed lines) are illustrated as a Gaussian with central value of zero, which is true for all the land regions in our inversion, and standard deviation equal to the annual uncertainty in the prior. The red lines in **Figure 3-1** define the posterior flux solution from the 91-site CE network inversion described in Chapter 2. The posterior solution for the first perfect flux test (prior set to known adjustments) is shown in dark blue lines, and the solution for the second test (prior and uncertainty set to the known adjustments for this network) in cyan lines.

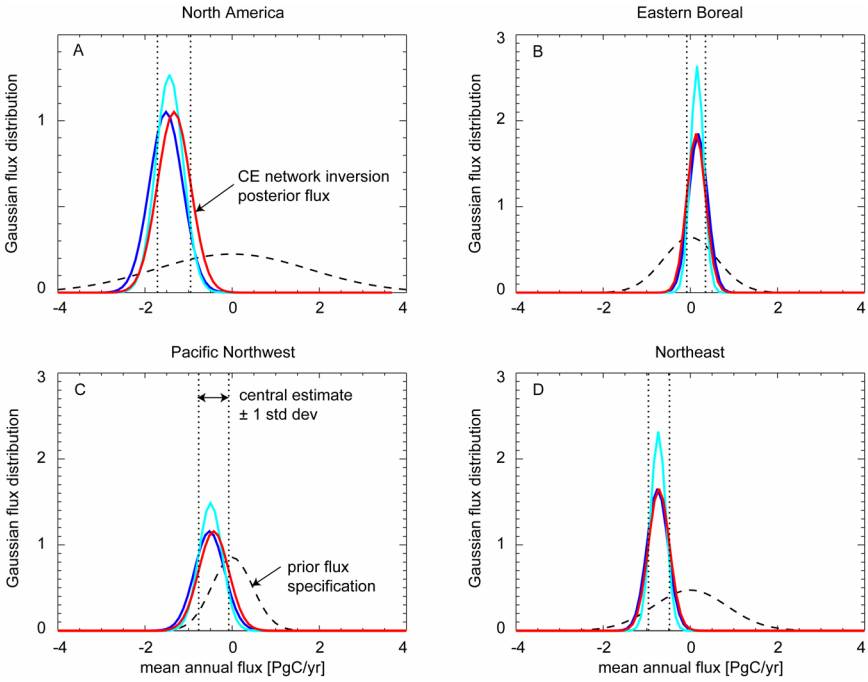

Figure **3-1**: Mean annual flux results from the perfect flux positive control experiments. Flux distributions are shown as Gaussians with central value equal to the mean annual flux and standard deviation equal to the mean annual uncertainty for A) North America, B) Eastern Boreal North America, C) Pacific Northwest Temperate North America, and D) Northeast Temperate North America. Dashed black lines are the prior flux specifications for the original CE network inversion. Red lines are the posterior solution for the CE network inversion. Vertical dotted lines define the central estimate ± uncertainty (1 standard deviation) of the CE network inversion solution. Dark blue lines are the posterior solution using the red solution adjustments as the prior flux specification. Cyan lines are the posterior solution using the red solution adjustments and posterior uncertainty as the prior specification.

The flux solutions for the perfect flux tests for the Eastern Boreal and Northeast regions are very nearly the same for the first test, and with the same central estimate but narrowed uncertainty range, as expected, for the second test (panels B and D in **Figure 3-1**). Recall from Chapter 2 that the posterior flux uncertainty is a function of the uncertainties in the data and the prior fluxes. Reducing the prior flux uncertainty in the second test requires that the posterior uncertainty for this test be no more than the prior uncertainty. Both the Pacific Northwest region, with no local observation sites, and the aggregated North American region show some offset, but all central estimates are within the posterior uncertainty of the original CE network inversion.

Globally, at the region-month level, the average difference in the posterior flux between the CE inversion and the first perfect flux test is -0.00054 Pg C yr^{-1} with a variance of 0.0894 (Pg C yr^{-1})^2. Also, as expected, the average monthly difference and variance for the second test are smaller at -0.00035 Pg C yr^{-1} and 0.0311(Pg C yr^{-1})^2, respectively.

For the perfect data tests, we first created two "observation" data sets; these are composites of the responses to the background fluxes at the sites in the 91-site CE network using either selected hours (daytime at continental, quasi-continuous sampling sites, standard times approximately weekly at discrete sampling sites) or using all hours. Each was used in an inversion variation in which the transport matrix matched the perfect observation data sets exactly. We expect that the inversions will return zero adjustments because the responses to the background fluxes in the transport matrix are matched in the observation data sets. We then created ten "noisy" versions of the first perfect data set by adjusting each monthly observation by a random amount within the monthly variability of the model samples. The variability of the model samples is, in general, greater than the monthly variability of the actual observations. The inversions with these noisy versions of the perfect observations should give some idea of how sensitive the inversion results are to larger than normal perturbations in the actual observations.

Figure 3-2 shows the results of these perfect and noisy data inversion tests, displayed in the same manner as the perfect flux experiments. The mean annual flux solutions for 2001-2003 are shown as Gaussian distributions using the flux estimate as the central value and the posterior uncertainty to represent 1 standard deviation. Results in **Figure 3-2** are for A) the aggregated North American region, B) Eastern Boreal, C) Pacific Northwest and D) Northeast sub-regions. The first perfect data inversion, with the selected hours observations, is shown by the red line, and the second perfect data inversion is shown in blue. Thin black lines represent the results of the ten noisy data inversions.

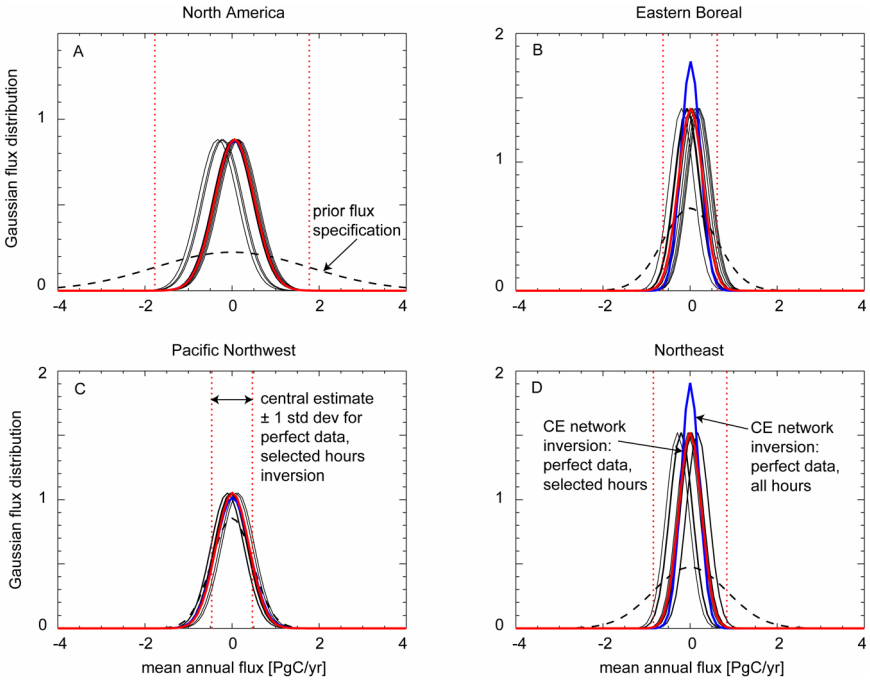

Figure **3-2**: Mean annual flux results from the perfect and noisy data positive control experiments. Flux distributions are shown as Gaussians with central value equal to the mean annual flux and standard deviation equal to the mean annual uncertainty for A) North America, B) Eastern Boreal North America, C) Pacific Northwest Temperate North America, and D) Northeast Temperate North America. Dashed black lines are the prior flux specifications for the original CE network inversion. Red lines are the posterior solution for a CE network inversion using as observations the site responses to the background fluxes for selected hours. Blue lines are for a similar inversion using as observations site responses for all hours. Thin black lines are the inversion results for ten inversions using perturbed versions of the perfect observations for the inversion shown with red lines. Vertical dotted lines define the central estimate ± uncertainty (1 standard deviation) of the perfect data inversion with red lines.

Both perfect data inversions replicate the central value of zero adjustment to the background fluxes, as expected, and the mean data residuals for these two perfect inversions are very small, -0.001 ppm. It is interesting to note that the choice of selected hours vs. all hours for the perfect data makes a difference in the Eastern Boreal and Northeast sub-regions, both of which have quasi-continuous observation sites, but no difference in the Pacific Northwest, with no observation sites, or the aggregated North American region as a whole. The smaller uncertainty in the all-hours perfect data inversion may be due to the greater number of hours

included in the monthly variability calculation at the discrete sampling sites. It is difficult to imagine that the monthly variability at the continental, continuous sites would be smaller when including the night-time hours when the accumulation of carbon dioxide concentrations in the surface layer is common at some of the continental sites.

Ten realizations of perturbed perfect data are probably not enough to draw clear conclusions, but it is interesting to see what appear to be two modes of the flux solutions for North America (panel A in **Figure 3-2**), one of which is distinctly a carbon sink where none is indicated in the annually neutral terrestrial background flux. None of the noisy data inversions, however, produces a central estimate outside 1 standard deviation from the first perfect data inversion test.

Overall, the method appears to be producing expected results in these controlled situations. We find no obvious instabilities.

3.3. Sensitivity to the Background Fluxes Used

To examine the influence of the background fluxes used in the control inversion, we use three variations of the inversion with the same CE network (**Table 3-1**), but with alternative terrestrial and fossil fuel background fluxes. For the hourly, annually varying SiB3 terrestrial background flux (Baker et al., 2007) we substitute the CASA climatology (Randerson et al., 1997) as used in the TransCom experiments. Both of these terrestrial fluxes are annually neutral at each land surface grid cell, but differ, especially in the northern latitudes, in the amplitude and timing of the seasonal cycle (Law et al., 2008). For the seasonally varying fossil fuel emissions (Erickson et al., 2008) used in the control inversion, we substitute the non-varying annual fossil emissions maps from which the seasonally varying emissions were created by approximating the

seasonal carbon dioxide emission cycle using a 2-harmonic Fourier series with coefficients as a function of latitude.

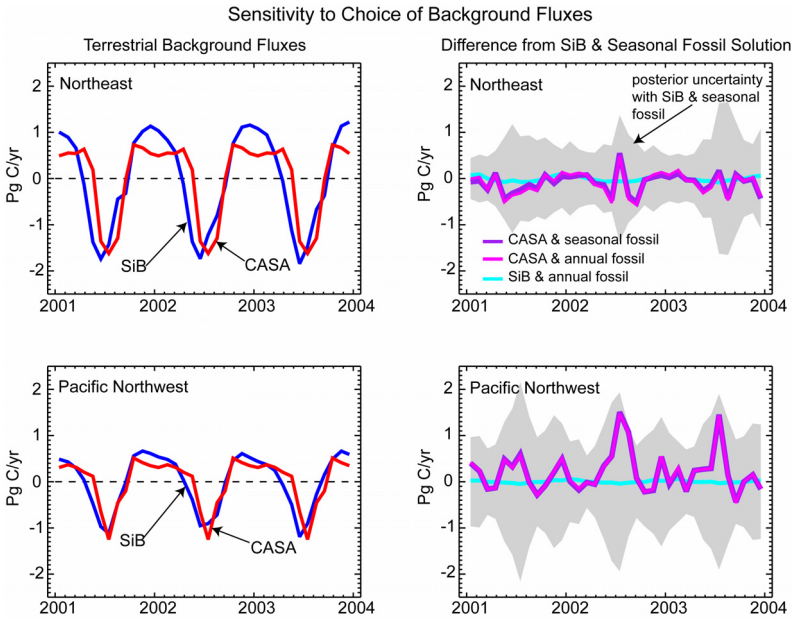

Figure **3-3**: Sensitivity of the inversion solution to alternative background fluxes for two regions in Temperate North America: Northeast (top) and Pacific Northwest (bottom). Left hand panels illustrate the two terrestrial background fluxes used. Right hand panels show the difference in the posterior monthly fluxes compared to the control inversion for three combinations of two terrestrial background fluxes and two fossil fuel emission maps. Shading indicates the posterior uncertainty of the control inversion.

Figure 3-3 illustrates the difference in posterior monthly fluxes using different combinations of the two terrestrial background fluxes and two fossil fuel emissions maps. The SiB3 terrestrial background flux is an hourly product with annual variability and the CASA terrestrial flux is a monthly mean climatology. In these regions and for the monthly time resolution of this inversion experiment, there is little evidence that the choice of terrestrial background flux is the primary determinant of the posterior flux solution. Both background fluxes can be corrected to a similar solution determined by the observations and the transport. This is not necessarily true in regions without nearby observational constraint.

The choice of fossil fuel emissions map, seasonal or annual, results in at most a 0.1 Pg C yr^{-1} difference in any monthly flux in the regions shown in **Figure 3-3**. The largest differences are in mid-summer and mid-winter in the northern mid-latitudes, with north central Europe having a 0.2 Pg C yr^{-1} difference in some months. The monthly differences are consistent with the PCTM results, in the mid-range for the three transport models tested in Gurney et al. (2005), in which a number of different amplitudes of seasonal variations were tested with seasonal mean inversions. There is no difference in the annual flux solutions, which is also consistent with the findings of Gurney et al. (2005) in comparing annual mean inversion results using seasonally-varying and non-varying fossil emissions maps. We did not, however, investigate the bias that may be present in our results due to the changing global distribution of fossil fuel emissions since the spatial distribution of the 1990's era maps used in our experiment.

3.4. Sensitivity to Magnitude of Prior Flux Uncertainty

The prior fluxes were chosen to specify no correction to the background fluxes, but with relatively large assigned uncertainty compared to previous studies (Chapter 2, Section 2.2.5). The prior flux uncertainties used are of very similar magnitude to those used in the TransCom Interannual Variability network sensitivity experiment (Gurney et al., 2008) for ocean regions and larger than the TransCom uncertainties for land regions. We experiment here with the magnitudes of these prior flux uncertainties. The TransCom convention is to use supplementary land use change priors in addition to the terrestrial background fluxes in some regions (e.g., Tropical Asia). We have not used these, instead choosing to add an additional biomass burning background flux. As an initial test there is an inversion variation (**Table 3-1**) with uncertainty magnitudes kept as close as possible to those used in Gurney et al. (2008) by distributing uncertainty from the larger TransCom aggregated regions (**Table A-1**) to the smaller regions of

our method based on the underlying flux pattern used to establish the response functions (Chapter 2, Section 2.2.3). In addition we computed inversion variations with the magnitude of the prior flux uncertainties from the control inversion multiplied by factors of 0.5, 2, 5, 10, and 50. With this inversion method it is common to see large month-to-month variations in the regional flux solutions when the prior uncertainties are loosened, even though the inversion calculation enforces global annual net flux totals. The monthly regional flux solution can become

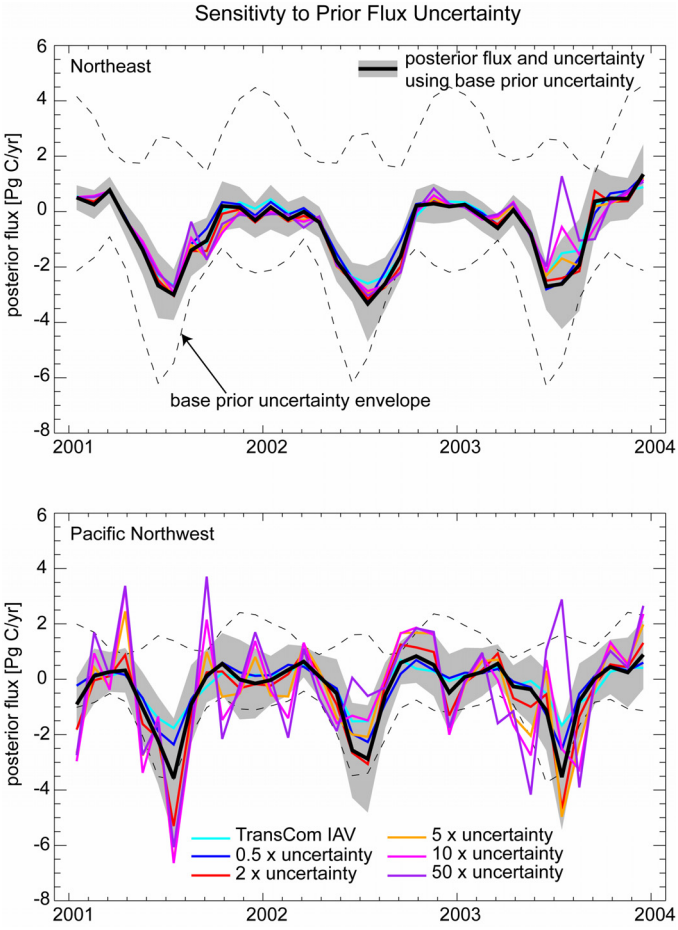

Figure **3-4**: Sensitivity of the inversion solution to the magnitude of prior flux uncertainties for two regions in Temperate North America: Northeast (top) and Pacific Northwest (bottom). Solid black line and shading represents the control inversion. Dashed lines define the prior uncertainty envelope for the control inversion.

biogeochemically unrealistic; tighter prior uncertainties have been used in the past to exclude these unreasonable solutions. This is especially problematic in regions of the world with sparse observations.

Figure 3-4 illustrates the results in two Temperate North American regions of loosening the prior flux uncertainties. The Northeast region has two local observation sites; the posterior uncertainty is markedly reduced compared to the prior uncertainty in the control inversion. In this region, multiplying the prior uncertainty by up to 5 does not change the result within the posterior uncertainty of the control inversion. The Pacific Northwest region is less well-constrained with no local observations. Noise in the posterior monthly fluxes is evident as the prior uncertainties are loosened. This effect is even more obvious in regions with very low observation density (e.g., regions in South America). Without local observations the tighter prior uncertainties are necessary to prevent unrealistic solutions. Prior flux uncertainties must be chosen carefully in light of the available data constraint.

3.5. Sensitivity to a Single Station

Another indication of problems with the solution method is sensitivity to outlier data points (Tarantola, 2005). In developing the base and enhanced networks in Chapter 2, we found one observation site which caused a large difference in the monthly regional flux solution. Here we show the inversion behavior that caused us to eliminate the site from consideration and try to find the cause of the behavior. Estevan Point is a discrete sampling site on the coast of British Columbia in the Environment Canada (EC) measurement program. When we included this site, the inversion solution exhibited anomalous fluxes in the Western Boreal North America region and compensating fluxes in the Northern and Eastern Boreal regions as shown in Panel A of **Figure 3.5**.

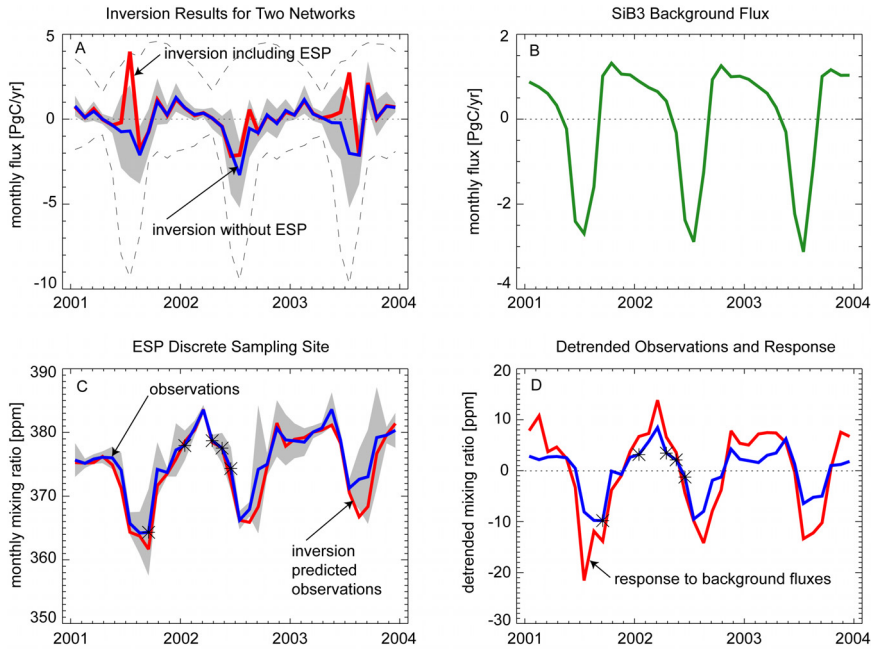

Figure **3-5**: Sensitivity of the Western Boreal North America region to the observation at Estevan Point (ESP). Panel A shows the Western Boreal monthly posterior fluxes and uncertainty for the control inversion (blue line and shading) and the posterior fluxes for an inversion including the ESP site (red line). Panel C shows the observations and monthly variability (blue line and shading) and the inversion-predicted observations for the inversion using the ESP site (red line). Data points marked with asterisks (*) in the observations are gap-filled. Panel B shows the SiB3 terrestrial background flux for the region. Panel D shows the response to the composite background fluxes sampled from the model at the ESP location (red line) and the observations (blue line); both time series in Panel D are detrended.

The red line in Panel D of **Figure 3-5** is the sum of the responses to the four background fluxes co-sampled from the model at the ESP location. The SiB3 terrestrial flux in Panel D is one of those background fluxes; the others are the response to the fossil fuel emissions, the biomass burning emissions and the air-sea flux. It is clear that the amplitude of the seasonal cycle of the responses to the background fluxes exceeds that of the observations. The drawdown in summer of 2001 is overestimated by more than 10 ppm. In order to fit the observations, the inversion requires an efflux in both 2001 and 2003 to fit the excessive summer drawdown. Estevan Point is

a coastal site, especially prone to representation errors in the transport models. Sampling at an offshore grid cell is often recommended in order to match the protocol of sampling marine air. In the PCTM model, the grid cell for the Estevan Point site is designated as sea; we also sampled at an adjacent land grid cell, but did not test the inland model sampling location in an inversion. We have not seen this behavior in other regions in North America, but it is evidence of the sensitivity to some observation locations.

The CE network used in the control inversion (**Table 3-1**) has two observation sites in the Northeast region in Temperate North America which are relatively close to each other: the flux tower sites at Harvard Forest in Massachusetts, USA (Urbanski et al., 2007) and Howland Forest in Maine, USA (Hollinger et al., 1999). In any effort to design a future network, it is useful to know how many observation sites are needed to characterize regions in the inversion. Here we compare (**Table 3-1**) monthly regional posterior fluxes for the Northeast Temperate North America region for the enhanced network inversion (with neither site), network N6 inversion (with the Howland Forest site, but not Harvard Forest), network N7 (with the Harvard Forest site, but not Howland Forest), and the control inversion (with both sites).

The results of this test are shown in **Figure 3-6**. The difference in the networks used in Panel A and Panel D of the figure is the addition of five flux tower sites, four of which are in North America. The inversion results including either one of the Howland Forest or Harvard Forest sites show similar reductions in uncertainty, but different monthly fluxes. The observed mixing ratio record at Harvard Forest exhibited one of the lowest growth rates of any of the observation sites used; this likely accounts for the greater growing season sink in the inversion results including the Harvard Forest site without the second site in the region. The results using both sites are similar to the results with only the Howland Forest site, but with subtle increases in the amplitude of the seasonal cycle and some additional reduction in posterior uncertainty. Two observation sites in the Northeast region, not sampled in the same model grid cell, appear to result

in a solution which accommodates the observations from both sites, countering in some aspects the problem of choosing one site to be representative of the entire region.

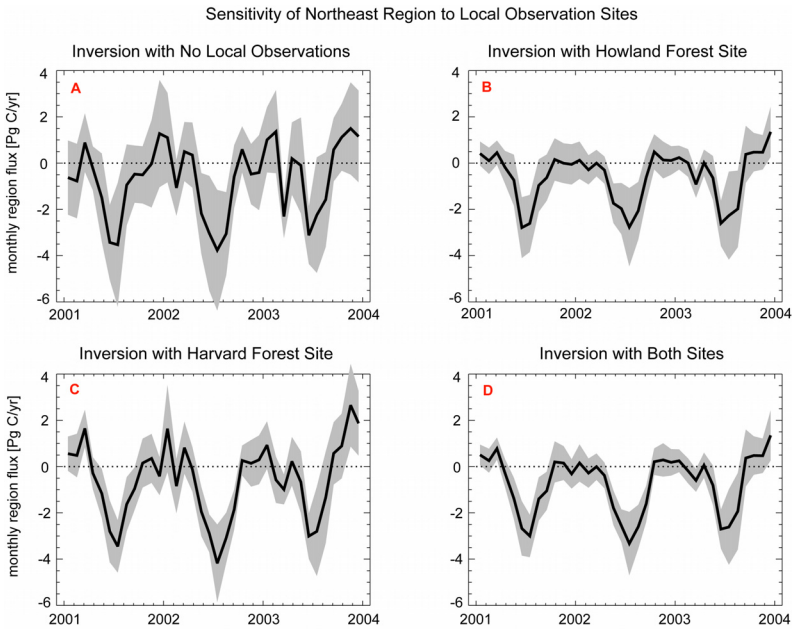

Figure **3-6**: Monthly posterior inversion fluxes (solid lines) and uncertainty (shading) for the Northeast region of Temperate North America for inversion variations for four networks: (A) the E (enhanced) network with no observation sites in the Northeast region; (B) network N6, the CE network without Harvard Forest; (C) network N7, the CE network without Howland Forest; (D) the CE network inversion with both sites.

3.6. Chi Square Test and Degrees of Freedom

With incomplete knowledge of the errors and number of degrees of freedom in the observation data, the principal consequence is that the posterior flux uncertainties, derived in part from the observation errors, will be flawed (Tarantola, 2005). Taratola (2005) recommends that

an analysis of the appropriateness of the prior uncertainties should include an evaluation of the χ^2 statistic, which is twice the cost function (equation 2.2 and repeated here):

$$J = (\mathbf{c_{obs}} - \mathbf{c_{fwd}} - \mathbf{Hx})^T \mathbf{R}^{-1} (\mathbf{c_{obs}} - \mathbf{c_{fwd}} - \mathbf{Hx}) + (\mathbf{x} - \mathbf{x_0})^T \mathbf{P}_0^{-1} (\mathbf{x} - \mathbf{x_0}) \quad (3.1)$$

As this is an implementation of a least squares minimization, this cost function at its minimum should follow the chi square probability density distribution, with the number of observations as the degrees of freedom. Previous researchers have used a normalized χ^2, which is divided by the number of observations and should, therefore, be close to 1 (e.g., Peylin et al., 2002; Rödenbeck et al., 2003a). The rationale is that a value >1 indicates that the residuals are larger than the prior uncertainties (observations and/or flux), suggesting that the prior uncertainties should be increased. In these other studies this diagnostic is used to suggest supplementing the data uncertainties to account for the model-data mismatch and "tune" the inversion. The normalized χ^2 for our control inversion is not larger than 1, but rather is ~0.50; however, equating of the number of observations with the degrees of freedom is problematic as we know that the observations are not independent in time or space. The degrees of freedom must be equal to or smaller than the number of observations, but we have not yet determined what it should be based on our use of co-located observations and the autocorrelation of the time series. Previous research (e.g., Peylin et al., 2002; Rödenbeck et al., 2003a) has concentrated analysis on the data part of the cost function for each observation site (called the χ^2 by station) for two main reasons. It is possible to analyze the data residuals directly as the analytic inversion solution delivers a set of predicted observations consistent with the flux solution. The second and more important reason is the presumption that the major portion of the normalized χ^2 is due to the data portion of the cost function (equation 3.1). For our control inversion, the mean χ^2 by station is 0.36. In inversions using a smoothed data product for observations (e.g. Baker et al., 2006), the

data uncertainties are often based on the residuals between the actual observations and the smoothed data used in the inversion, with an additional misfit error added for some observation sites to account for poor fit of the modeled data to the observations. The background terrestrial flux, in particular, may not be a good match for the amplitude and/or timing the seasonal cycle at some observation sites; in addition, at coastal sites, the resolution of the transport may not be fine enough to represent local conditions. By adding model-data misfit error to acknowledge the transport and representation error, we effectively de-weight the observations, putting correspondingly more weight on the background fluxes. In our experiment, the data uncertainties represent the monthly variability of the observations directly. This should be a conservative error estimate; however, we are hesitant to reduce the data uncertainties as might be indicated by the value <1.

To remove one of the issues about independence of the observations we calculate an inversion variation (**Table 3-1**) removing co-located surface observations (network N8 in **Table A-4**; Observation Site-Network Cross Reference in **Table A-9**). We remove the seven discrete sampling sites that are co-located with quasi-continuous sampling sites, the two lower levels at each of the two tall towers, and two discrete sampling sites that share the same grid box as other discrete samples. The normalized χ^2 for this inversion variation is 0.48 with the mean χ^2 by station 0.32, very similar to the control inversion. As a rudimentary test of adding model data mismatch error, we devised two different schemes for adding uncertainty to the monthly observations: level 1 with a moderate amount of mismatch error; level 2 with a larger amount. We assign observation sites to classes based on location (e.g. remote, coastal or continental) and model fit, and increase monthly data uncertainty based on the class. For example, remote, quasi-continuous sites are assigned 0.50 ppm additional uncertainty in the level 1 scheme and 0.75 ppm in the level 2 scheme. The largest mismatch errors are assigned to observation sites whose responses to background fluxes are very poor fits to the observations; for these sites, the level 1

added uncertainty is 5.00 ppm and the level 2 added uncertainty is 7.50 ppm. Both the classification of sites and the specific model-data mismatch error assignments should be considered preliminary. We added two more inversion variations with the same N8 network using the observation data uncertainties supplemented with this mismatch error. The inversion variations with added data uncertainties yielded normalized χ^2 mean values of 0.20 (level 1, moderate added error) and 0.13 (level 2, more added error).

Table 3-2 shows the inversion results for the aggregated global regions (**Table A-1**) for the control inversion, the no co-located sites inversion, and the two mismatch error inversions. In comparing the control inversion and the N8 network inversion before adding mismatch error, there is some re-distribution of the sources and sinks, but no changes outside the range of the posterior uncertainties of the control inversion. The balance of mean annual sink between Boreal Asia and Europe shifts between the inversions for these two networks. Within North America there is a shift of sink from the Pacific Northwest to the Western Boreal region, also within the posterior uncertainties of the control inversion. A comparison of the posterior annual flux covariance matrices for these two inversions shows no change in the covariance structure and only small changes in the variances as seen in **Table 3-2**. A recommendation for future inversion design is to use only one observation time series from each location; this also avoids one of the issues in determining the number of degrees of freedom represented in the observation data.

The main difference that results from adding mismatch error to the N8 network observations is further shift of the distribution of the mean annual sink between the Boreal regions, now involving Boreal North America as well as Boreal Asia and Europe. The increased posterior uncertainties will mathematically result from the increased the data uncertainties. More investigation is necessary to determine the appropriate use of the mismatch errors in this solution method.

Table **3-2**: Mean annual 2001-2003 inversion results (flux ± uncertainty) for aggregated land and ocean regions for the 91-site CE network control inversion, the 78-site N8 network inversion (no co-located measurement programs), the N8 network with moderate added model-data mismatch error (Mismatch 1), and the N8 network with larger added model-data mismatch error (Mismatch 2). All units are Pg C yr^{-1}.

Region	Control Inversion	Network N8 Inversion	Mismatch 1 Inversion	Mismatch 2 Inversion
Aggregated Land Regions				
Boreal North America	0.00 ± 0.32	-0.04 ± 0.34	-0.25 ± 0.39	-0.38 ± 0.47
Temperate North America	-1.34 ± 0.40	-1.17 ± 0.42	-1.15 ± 0.46	-1.12 ± 0.51
Tropical America	0.22 ± 0.79	0.16 ± 0.81	0.08 ± 0.84	0.01 ± 0.88
Temperate South America	0.68 ± 0.87	0.63 ± 0.89	0.46 ± 0.94	0.27 ± 0.98
Northern Africa	0.86 ± 0.65	0.98 ± 0.67	0.81 ± 0.74	0.65 ± 0.82
Southern Africa	-0.60 ± 0.67	-0.62 ± 0.68	-0.53 ± 0.73	-0.43 ± 0.77
Boreal Asia	-0.48 ± 0.49	-0.28 ± 0.52	-0.13 ± 0.64	-0.14 ± 0.81
Temperate Asia	0.05 ± 0.33	-0.22 ± 0.34	-0.33 ± 0.43	-0.36 ± 0.53
Tropical Asia	-0.13 ± 0.31	-0.11 ± 0.31	-0.03 ± 0.35	0.04 ± 0.39
Australia	-0.14 ± 0.38	-0.12 ± 0.38	-0.06 ± 0.41	0.02 ± 0.43
Europe	-0.35 ± 0.45	-0.58 ± 0.47	-0.28 ± 0.57	-0.05 ± 0.67
Ocean Regions				
North Pacific Ocean	-0.28 ± 0.23	-0.27 ± 0.24	-0.33 ± 0.29	-0.35 ± 0.36
Tropical West Pacific Ocean	0.03 ± 0.09	0.05 ± 0.09	0.07 ± 0.09	0.08 ± 0.10
Tropical East Pacific Ocean	0.18 ± 0.29	0.24 ± 0.30	0.30 ± 0.32	0.35 ± 0.33
South Pacific Ocean	-0.33 ± 0.23	-0.32 ± 0.24	-0.29 ± 0.25	-0.27 ± 0.26
Northern Ocean	-0.30 ± 0.16	-0.26 ± 0.16	-0.30 ± 0.19	-0.29 ± 0.21
North Atlantic Ocean	-0.15 ± 0.17	-0.14 ± 0.18	-0.13 ± 0.21	-0.14 ± 0.24
Tropical Atlantic Ocean	0.11 ± 0.11	0.10 ± 0.11	0.11 ± 0.11	0.12 ± 0.11
South Atlantic Ocean	-0.15 ± 0.13	-0.14 ± 0.13	-0.14 ± 0.13	-0.13 ± 0.13
Southern Ocean	-0.38 ± 0.20	-0.38 ± 0.20	-0.37 ± 0.23	-0.36 ± 0.25
Tropical Indian Ocean	0.17 ± 0.16	0.17 ± 0.16	0.17 ± 0.16	0.17 ± 0.17
South Indian Ocean	-0.27 ± 0.28	-0.28 ± 0.28	-0.31 ± 0.32	-0.34 ± 0.35

Temporal autocorrelation must also be considered. For a very rudimentary analysis from the time series perspective, we note that the autocorrelation lag for the monthly observations for most temperate latitude sites is ~3 months reflecting the terrestrial seasonal cycle. Using as a rough approximation that the number of degrees of freedom should be reduced by a factor of 3 would make the normalized χ^2 value larger than the goal of 1. This suggests supplementing the data uncertainties as we have done in the misfit tests. This issue deserves further research using the statistical tools appropriate for analyzing time series.

3.7. Examination of the Data Residuals

The monthly data residuals for 2001-2003 for the 91 sites in the control inversion were pooled for overall analysis. **Figure 3-7** shows the histogram, autocorrelation function, partial autocorrelation function, and q-q plot for the data residuals (Shumway and Stoffer, 2006). It is clear from the histogram and the q-q plot that the distribution of the residuals is more sharply peaked than a Gaussian distribution. The mean residual value for this control inversion is -0.128 ppm; other inversion variations show this consistent bias. There is little change in the form of the distribution of the data residuals for the N8 Network, Mismatch 1 and Mismatch 2 inversion variations, although the mean residual values are smaller: -0.121 ppm, -0.095 ppm, and -0.056 ppm, respectively. As the mismatch error is increased, effectively de-weighting the observation data, a seasonal pattern in the residuals becomes clearer in the autocorrelation and partial autocorrelation functions. We believe this is due to the relative increase in the weight of the background terrestrial biosphere flux in the inversion.

Using monthly means of observation data directly rather than a smoothed data product, we expect that there will be some outliers in the residuals for sites that may not be fit well when the background fluxes and/or transport are not sufficiently finely resolved. Neither the autocorrelation function nor the partial autocorrelation function shows significant violation of our assumptions about the errors in the data, although there is some cyclical residual structure in the partial autocorrelation. Using an autocovariance model, appropriate for time series, in the construction of the prior data covariance matrix $\mathbf{R}$ would be a good first step in addressing this issue. Examination of residuals at a site level might also guide future choices of sites to be used in the inversion.

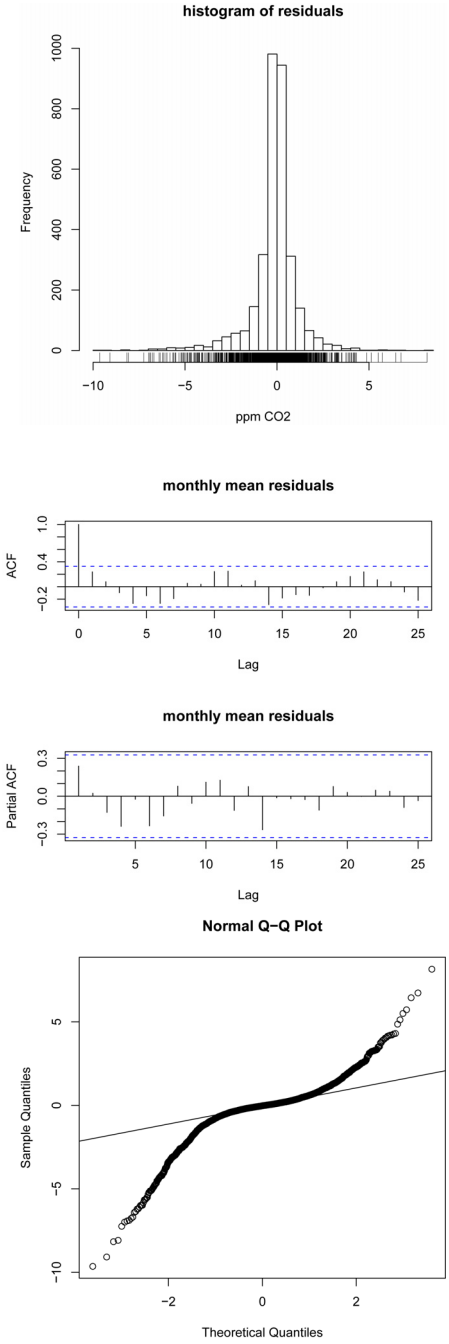

Figure **3-7**: Histogram, autocorrelation function, partial autocorrelation function, and Q-Q plot for the pooled data residuals for the 91 sites in the CE network control inversion.

For a different perspective on the quality of the inversion-predicted observations, **Figure 3-8** shows a Taylor plot for the control inversion (Taylor, 2001). The angular coordinate of each point on this plot is the correlation of site time series from the composite background fluxes (hollow symbols) or the inversion-predicted time series (filled symbols) with the monthly observed time series. The radial coordinate is the mean pre-inversion or post-inversion standard deviation of the monthly time series normalized by the standard deviation of the observed time series. The correlation is a measure of a match to the seasonal cycle of the observations and the normalized standard deviation is a measure of the match to the amplitude of the observations. This inversion method is designed to deliver the posterior fluxes that best explain the observations, so we expect that the post-inversion data points will cluster near correlation 1.00 and normalized standard deviation of 1.0. In short, this is further confirmation that the inversion has worked.

A few of the observation sites with less-than ideal post-inversion results are called out in **Figure 3-8** as examples. The observation site with the worst pre-inversion model fit to the data, the Tapajos flux tower site (TPJ), improves by this measure from correlation, normalized standard deviation pair of (0.55, 2.40) to (0.68, 1.46), but is still the worst fit site post-inversion. The Southern Great Plains flux tower (SGP) also shows improvement from a poor model fit. The St. David's Head, Bermuda (BME) discrete sampling site likely cannot be fit well by the inversion due to conflict with the neighboring BMW discrete sampling site. BME is one of the sites omitted from the N8 no co-located surface sites network inversion (**Table 3-1**). Both BME and BMW sites are sampled from the same model grid cell, but not at the same times due to our co-sampling protocol.

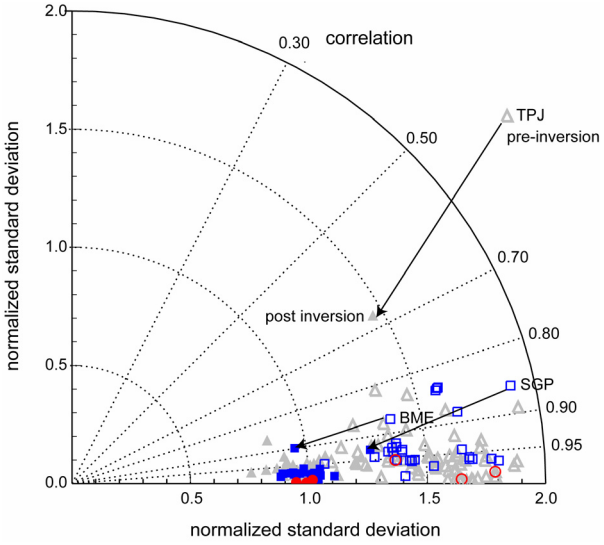

Figure **3-8**: Taylor diagram illustrating the pre-inversion (hollow symbols) composite background mixing ratios and post-inversion (filled symbols) predicted mixing ratios by station compared by correlation and normalized standard deviation to the observations at the 91 sites in the control inversion. Blue square symbols indicate North American sites; red circle symbols indicate NOAA ESRL observatories outside of North America. Arrows indicate the movement within this measurement context for a few example sites (Tapajos flux tower, Southern Great Plains flux tower, and the St. David's Head, Bermuda discrete sampling site).

We also calculated Taylor skill scores (Taylor, 2001) for each inversion as

$$S = \frac{4\left(1+R\right)^{p}}{\left(\hat{\sigma}_n + \frac{1}{\hat{\sigma}_n}\right)^{2}\left(1+R_0\right)^{p}} \tag{3-2}$$

where S is a skill score, R is the correlation of the of the predicted observations with the observations, $\hat{\sigma}_n$ is the normalized standard deviation of the predicted observations relative to the observations, R_0 is the maximum possible correlation (in this case, 1.00), and $p = 4$ at the suggestion in Taylor (2001) to impose a larger penalty on poor correlation. Scores tend to approach 1.00, as both the correlation and normalized standard deviation approach 1.00. One reason to calculate the score as an overall measure of the inversion's ability to fit the observation data is to determine if any one set of background fluxes produces a "better" score. We find that

there are very high scores (e.g., a mean of 0.94 and median 0.96 for the control inversion) and no discernible difference in the scores for different combinations of the SiB3 and CASA background terrestrial fluxes and the annual and seasonal fossil fuel emissions maps. We cannot conclude, using this score, which combination of background fluxes could be perceived as better to use with this inversion method.

3.8. Conclusions about Methods Choices

We find no excessive reliance on the background fluxes or magnitudes of the prior uncertainties for regions with sites in the network of observations. Although the inversion variations using the two alternative terrestrial background fluxes yield slightly different results, we find no reason to reject either. The seasonal fossil fuel emissions background flux likely reduces the fluxes that are aliased in the terrestrial flux corrections when using emissions with only annual variability. Using a fossil emissions map with current spatial distributions should be a priority for future experiments. The posterior fluxes reported here for Temperate Asia, for example, may not be an accurate representation of terrestrial biological activity.

Co-located surface observation sites should not be included in future inversions, as they do not necessarily provide any additional information to the inversion and confuse the issue of the data independence. This is an aspect of the experiment design which we would change in the future. Having more than one observation site in a region (when not co-located) does appear to bring more information to the problem, and can improve the issue of choosing observations to be representative of entire regions.

During the course of this experiment we have calculated many inversion variations with different networks and different background fluxes. **Figure 3-9** illustrates the partitioning of the 2001-2003 mean annual flux between several pairs of aggregated regions and sub-regions of 20

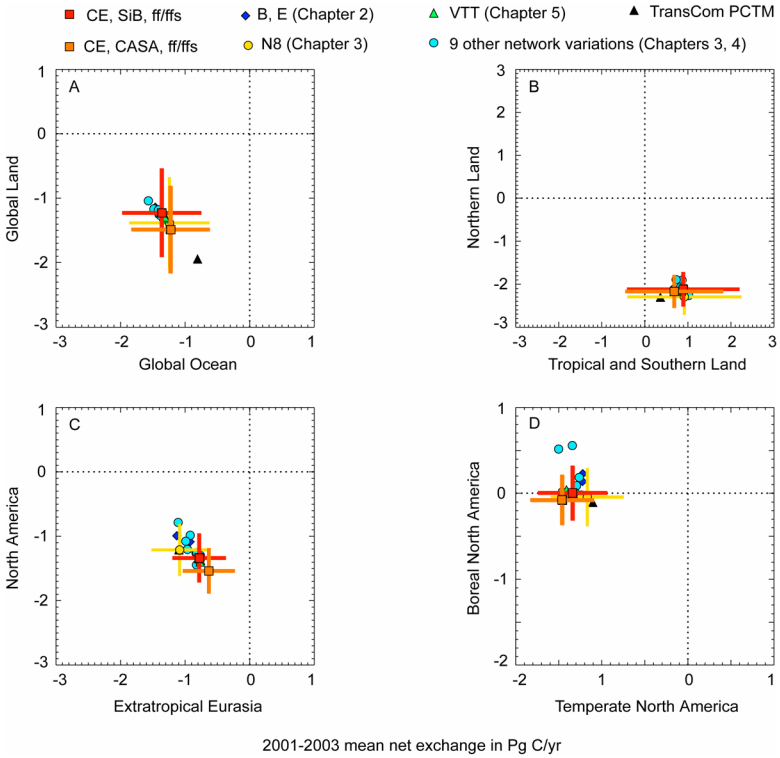

Figure **3-9**: Comparison of partitioning of mean annual fluxes for 2001-2003. Symbol shapes and colors indicate the inversion variations: CE network with SiB, annual and seasonal fossil (red squares with uncertainty); CE network with CASA, annual and seasonal fossil (orange squares with uncertainty); N8 network (yellow circles with uncertainty); B and E networks from Chapter 2 (blue diamonds); VTT networks from Chapter 5 (green triangles); 9 different network variations from Chapters 3 and 4 (cyan circles); TransCom PCTM (black diamond).

inversion variations, some of which will be discussed in Chapters 4 and 5. With the exception of the partitioning between Boreal and Temperate North America of two inversion variations (Panel D in **Figure 3-9**), the results are tightly clustered. The two inversion variations outside the main cluster in Panel D use networks with no continental, continuous observing sites in North America or globally. These continental sites are a key part of our experimental strategy and the primary reason for conducting the experiment. The tight clustering gives a boundary for the real (as opposed to the calculated) uncertainty of the posterior flux estimates for our experiment. The TransCom PCTM results for the same time period from the submission to the North American

Carbon Program Interim Synthesis is also shown (black triangle in **Figure 3-9**). The TransCom PCTM result is based on a different network, with fewer continental observation sites, which may account for the difference in the Global Land/Global Ocean partitioning in Panel A.

We have used a method utilizing a least squares analytic solution which may not be entirely appropriate. This method has been used extensively in the past, with each research group contributing to addressing the shortfalls. There are several assumptions still outstanding including the independence of the observations and prior errors. Future effort could be directed toward incorporating changes into the analytical method, or toward implementing a more general solution. The main consequence of using the method in spite of the shortfalls is that the posterior uncertainties are likely underestimated. The method still remains as a solid first approach to the problem. There is a future role of this method in the inversion toolkit for testing new observation sites and types or for producing flux estimates or boundary conditions for other inversion methods.

Chapter 4

Sensitivity to Network Choices

4.1. Introduction

In this chapter network design choices are explored. Results of two series of sensitivity tests are in Section 4.2. First is a series of inversions testing the sampling for selected hours of the day at continental observation sites and the co-sampling of model output for the same hours, as described in the methods in Chapter 2, Section 2.2.8. The second series tests the sensitivity of inversion results to the inclusion/exclusion of different classes of observation sites. The inversions discussed in this chapter are summarized in **Table 4-1**. The discussion is in Section 4.3. Conclusions and recommendations for future network design are in Section 4.4. Observation Site-Network cross references for these inversions are in **Table A-10**.

Table **4-1**: Inversion variations discussed in Chapter 4. The observation time series at continental, continuous observation sites are always sampled for selected hours of the day.

Inversion	Network	# of Sites	Description
Control Inversion	CE	91	Continental Extension Network inversion described in Chapter 2; model output co-sampled
Sensitivity to Model Sampling Protocol			
CE, all hours	CE	91	Model output sampled for all hours
CE, default hours	CE	91	Model output sampled for default hours
No continental continuous	T1	69	All continental continuous sites excluded; model output co-sampled
No continuous, all hours	T1	69	Model output is sampled for all hours; all continuous continental sites excluded
No continuous, default hours	T1	69	Model output sampled for default hours; all continental continuous sites excluded
Sensitivity to Inclusion/Exclusion of Classes of Observations			
No CAR, WPO	N1	73	Briggsdale, CO aircraft profile and Western Pacific Ocean high altitude aircraft transit excluded
POC, not WPO	N2	91	Pacific Ocean Cruise surface observations used in place of Western Pacific Ocean transit
No High Elevation	N3	86	European elevated surface sites excluded
No MNM, RYO, YON	N4	88	Minamitorishima, Ryori, Yonagunijima sites excluded

4.2. Results

4.2.1. Selected Hours Sampling at Continental Observing Sites

As described in Chapter 2, only mid-day hours (12-16 LST) of the observation time series are included in the calculation of monthly means for continental, quasi-continuous sites. The transport model output is co-sampled for the same hours. The selected hours sampling is based on two assumptions. During the mid-day hours at surface observing site, the continental boundary layer is well-mixed and we assume the carbon dioxide mixing ratio is representative of a broad area. We aim to maximize a single site as representative of a region used in the inversion. For elevated continental sites, the night-time hours are chosen instead to sample the residual well-mixed continental boundary layer from the previous afternoon. A second assumption is specific to the transport model and meteorological driver data, and asks the question: Is there a time of day when the model characterization of the continental boundary layer is most like the observations? Inversions using smoothed data products (e.g., Baker et al., 2006) and observing sites that sample for background conditions have used a model sampling strategy of averaging all hours of model output when calculating monthly responses. We compare the model output and observations using all hours of the day to the model and observation sampling protocol used in Chapter 2, using only the mid-day hours (continental surface sites) and night-time hours (elevated sites). **Figure 4-1** illustrates the differences in observed and modeled mixing ratios when using all hours compared to the co-sampled selected hours used for four continuous sites in the control network.

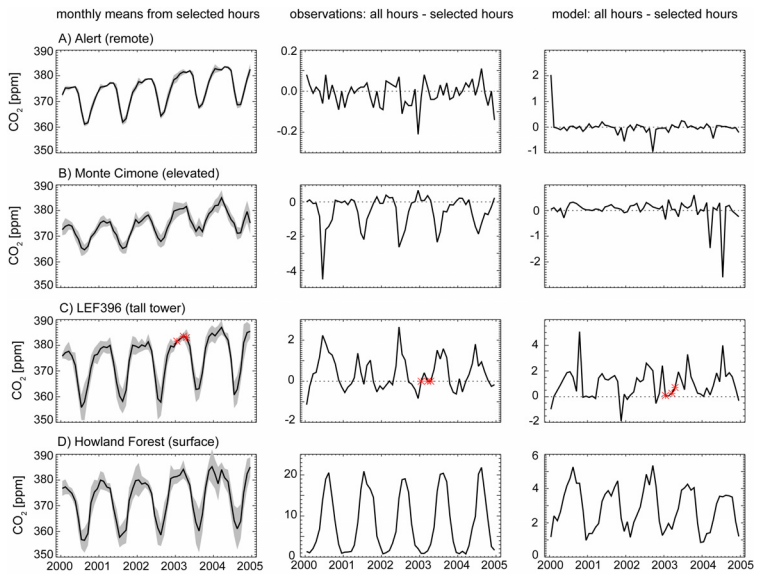

Figure **4-1**: Monthly mean mixing ratio and variability (shading), observation time series difference between monthly means from all hours and selected hours and modeled time series difference between monthly means from all hours and selected hours for row A: remote site with mid-day selection, B: elevated continental site with night-time selection, C: tall tower with mid-day selection and D: flux tower with mid-day selection. All units are ppm CO_2. Data points marked with asterisks (*) indicate gap-filled data.

At the remote site (row A in **Figure 4-1**), the difference between the monthly mean mixing ratios calculated for all hours compared to mid-day hours for either the observed time series or modeled mixing ratio time series is small (~-0.01 ppm for both observations and model samples). The elevated site (row B in **Figure 4-1**) shows a seasonal cycle in the difference between the monthly mean observed mixing ratio for all hours vs. the selected hours (annual mean difference -0.55 ppm), possibly corresponding to summer daytime photosynthesis not captured in the night-time sampling. The difference in the modeled mixing ratios (interpolated between mid-points of layers 4 and 5 above the model surface) is smaller, ~0.02ppm. At the tall tower site (row C in **Figure 4-1**) the annual mean difference in monthly means of mid-day vs. all hours observations in 0.38 ppm with maxima of 1-2 ppm in the summer months. The model is

sampled midway between model layers 2 and 3 at this site; the annual mean of the modeled differences is 0.99 ppm with essentially no difference in the wintertime months. The rationale for choosing mid-day hours at surface, continental continuous observation sites is most clearly illustrated at the flux tower site (row D in **Figure 4-1**). At this site the night-time buildup of carbon dioxide in the stable nocturnal boundary layer results in a mean difference between daytime hours and all hours of 8.25 ppm with values as large as 20 ppm in the summer. The difference in the modeled data is much smaller with summertime maxima of only 5 ppm. Choosing mid-day hours to compare observed and modeled mixing ratios excludes the nocturnal conditions which we assume to be poorly simulated by the model. We performed inversion variations with the monthly model samples created from all hours and the default sampling (Chapter 2.2.8.) to understand the consequences of choosing a model sampling protocol that does not match the selected hours used for the observations.

Inversion results testing different time-of-day model sampling protocols for continental quasi-continuous observation sites are shown in **Figure 4-2** for selected North American and adjacent land and ocean regions. The black bar and shading represents the posterior flux and uncertainty for the control inversion using the 91-site continental extension (CE) network. Symbols indicate the results of the five inversion variations related to model sampling protocol testing (**Table 4-1**). In all variations the observations at continuous sites included in the inversions are sub-sampled for selected times-of-day (mid-day for surface and tall tower sites, mid-night for elevated sites). The CE, all hours inversion (up-pointing triangles in **Figure 4-2**) is the result for the CE network with the model samples including all hours. This is the simplest model sampling strategy, and has been used in the past for inversions that use data products such as GLOBALVIEW (GLOBALVIEW-CO2, 2007) in place of observations (e.g., Baker et al., 2006). The CE, default hours inversion (diamonds) is the results for this same CE network where

model samples are chosen from default sampling protocol; samples from the selected times-of-day from every day are used.

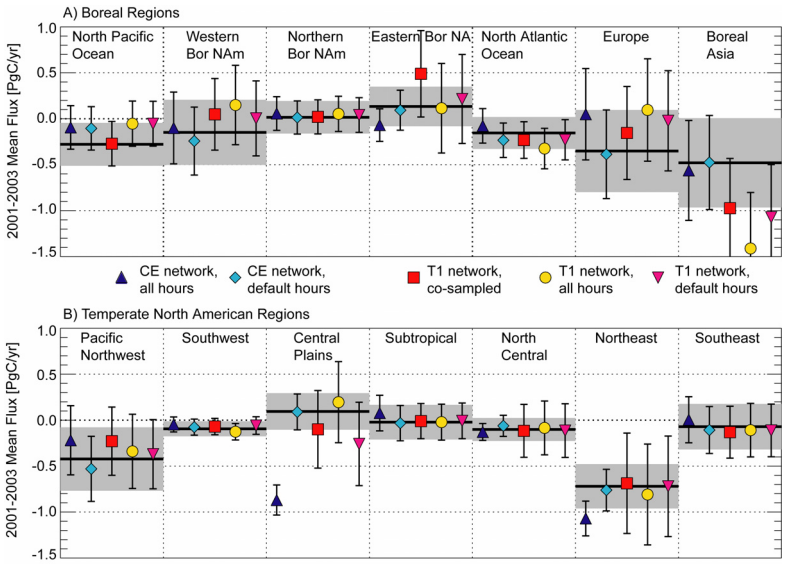

Figure **4-2**: Mean annual fluxes and uncertainties for 2001-2003 for A) Boreal and B) Temperate North American regions showing impact of different choices for time-of-day model sampling. Black line and shading represents the result of the continental extension (CE) network inversion with sub-sampling of model output at continental sites. All inversion results shown use sub-sampling of observations at continental sites. Inversion results shown are: CE network with model output sampling for all hours (up-pointing triangles); CE network with model sampling using default hours (diamonds); T1 network (69 sites with no continental, quasi-continuous sites) with co-sampling of model output at remaining sites (squares); T1 network with model sampling for all hours (circles); T1 network with default model sampling (down-pointing triangle). All units are Pg C yr^{-1}.

This represents the next step up in complexity from the all-hours sampling, but still simpler than co-sampling the same hours as the observations, which is the strategy used in the control inversion. If the model sampling strategy made no difference, we would expect the up-pointing triangles and diamonds to match the control inversion results. This is not the case, especially for the all-hours model sampling strategy in the regions of North America with the flux towers (Eastern Boreal, Central Plains, and Northeast) and Europe. At the remote Alert

observation site in the Northern Boreal region, the time-of-day selection made little difference for either model output or the observations (**Figure 4-1**). In **Figure 4-2** there is little difference in the inversion result as well. Results for the CE, default hours inversion are closer to the control inversion; if there were no gaps in the observations at the selected times-of-day, the default model sampling would be the same as co-sampling for the continuous sampling sites.

The three remaining inversions shown in **Figure 4-2** are inversions using the 69-site T1 network with the same co-sampling strategy used in the control inversion (squares), T1 network all hours model sampling (circles) and T1 network default model sampling (down-pointing triangles). The T1 co-sampling inversion (squares) results illustrate, by their omission, the role of the continuous sites in determining the uncertainties in the regions containing these observing sites in North America; for example, see the increase in uncertainties for the T1 co-sampled inversion (squares in **Figure 4-2**) compared to the control inversion for Eastern Boreal North America, Central Plains, North Central, and Northeast regions. In some regions the flux solution is also different, for example in the Eastern Boreal region where the Northern Old Black Spruce flux tower and Fraserdale observing sites in the control network both reduce the uncertainty and change the magnitude of the regional flux. If the model sampling protocol made no difference for the discrete sampling sites, we would expect the three solutions for this 69-site T1 network to be the same. Discrete sampling is done at times chosen for sampling of background atmospheric conditions. Neither the all hours model sampling nor the default sampling T1 inversions agree with the T1 co-sampled inversion in some regions. The difference in the T1 all hours inversion and the T1 co-sampled inversion suggests that, when using actual observations rather than a smoothed data product, it is not advisable to create monthly mean model output samples from all hours. The default hours sampling difference is not surprising; the model output for discrete sites is sampled on five fixed days in each month without regard to the meteorological conditions of the transport. These are not likely to be the days and hours that the discrete observation sampling

was done. The large differences in flux solutions for Europe and Boreal Asia across the inversions in **Figure 4-2** are further evidence of the effect of sparse observations in North Central Europe and Boreal Asia; the posterior uncertainties are large compared to the other regions shown for the control inversion. These uncertainties are unlikely to be reduced without additional observation sites in North Central Europe and Boreal Asia.

4.2.2. Sensitivity to Classes of Observations

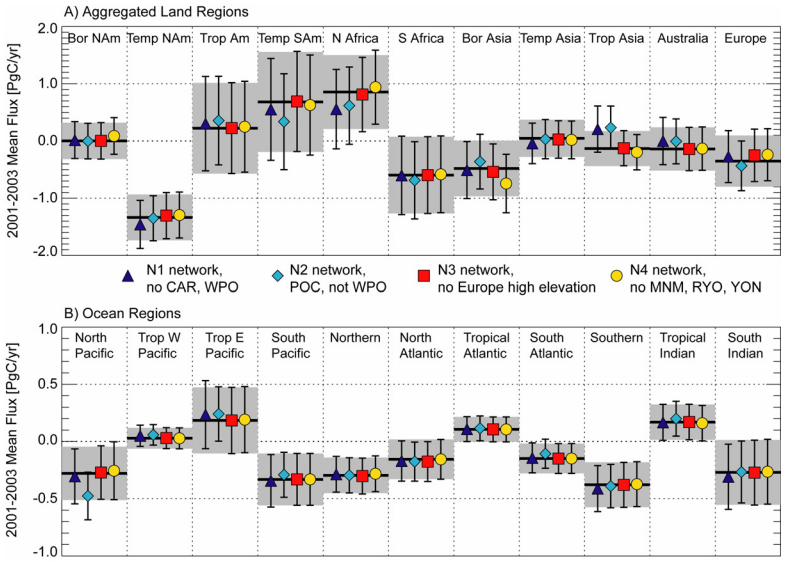

Figure **4-3**: Mean annual fluxes and uncertainties for 2001-2003 for A) aggregated land regions and B) ocean regions for inversions illustrating inclusion/exclusion of classes of observations. Black line and shading represent the result of the CE network inversion. Inversion results shown: N1 network with no aircraft observations (triangles); N2 network with Pacific Ocean surface observations used in place of high altitude Pacific Ocean transit observations (diamonds); N3 network with no European elevated surface sites (squares); N4 network omitting 3 east Asian sites. All units are Pg C yr^{-1}.

Figure 4-3 shows sensitivity of the continent- and ocean basin-scale inversion results to different classes of observations. Chosen here are several classes of observations that are challenging to use in inversions for a variety of reasons: aircraft profiles, aircraft high altitude

samples, high elevation continental sites, and sites that are not modeled well by the background fluxes. The inversion results indicated by triangles in **Figure 4-3** are from the 73-site N1 network that eliminates all aircraft measurements, namely the Briggsdale, Colorado (CAR) vertical profile and the high altitude western Pacific Japan to Australia transit (WPO). The North American results change little, but there is a loss of constraint and shift in flux for Temperate Asia, Australia, and especially Tropical Asia which is a net source and no longer neutral with this network. There are also compensating sinks introduced in the ill-constrained Temperate South America and North Africa. The Pacific Ocean ship of opportunity (POC) observations are an alternative set of observations in the Pacific Ocean basin with approximately the same latitudinal range as the WPO series. We substitute gap-filled POC observation time series for the WPO observations in the 91-site N2 network to test this different Pacific Ocean basin constraint; the results are shown as diamonds in **Figure 4-3**. These ocean surface observations are more commonly used in global atmospheric inversions, but were not included in our control inversion due to a lengthy observation gap in the middle of our chosen years of analysis. With the exception of the North Pacific, where there is an increased sink using the ocean surface observations, and a shift in fluxes between Boreal Asia and Europe, the inversion fluxes are similar to the N1 network (no-aircraft) inversion. Other contemporary inversions (e.g. Peters et al., 2007) find a larger sink in the North Pacific in agreement with the N2 network inversion. It is likely that the high altitude Western Pacific observations and the Pacific Ocean surface observations bring different information to the inversion; both could be used when the observations fill the criteria for inclusion.

Figure 4-3 also shows the results of the 86-site N3 network inversion which omits the mountain-top sites in Europe for which mid-night hours of observations are used (squares). The shift in fluxes between Boreal Asia and Europe, with the North Central sub-region of Europe most affected is consistent with what we reported in Chapter 2. The last inversion shown (circles)

is for the 88-site N4 network omitting the Japanese Meteorological Agency east Asian marine and coastal sites of Minamitorishima (MNM), Ryori (RYO), and Yonagunijima (YON). The inversion-predicted carbon dioxide time series at the RYO site, for example, is consistently lower than the observed time series with a mean 36-month residual of -3.1 ppm in the control inversion; this is the largest mean residual of all the sites used in this inversion, but within the monthly variability of the monthly mean observations. Omitting these east Asian sites increases the sink in Boreal Asia and Tropical Asia with offsetting reductions in sinks distributed among Europe and some of the less well-constrained aggregated regions. There is little change in the fluxes in the Pacific Basin, but there is an increase in uncertainty, particularly in the North Pacific.

4.3. Discussion

4.3.1. Selected Hours Sampling at Continental Observing Sites

The difference in monthly means of carbon dioxide observations calculated from all hours and those from mid-day hours can be on the order of 20 ppm in the summer at continental surface sites where a nocturnal stable layer buildup is common. While this is not true at all continental sites, similar nocturnal buildups represent local phenomena, and not the larger region to be represented in an atmospheric inversion. The buildup appears not to be modeled correctly in the transport used here. The modeled samples may reflect too much vertical mixing in the surface layer at night resulting in a much smaller vertical gradient than seen in the observations. If we are trying to match model representation to local observations, confining both model and observation sampling to mid-day hours seems to be a reasonable approach for continental surface sites. To a lesser extent, the same logic applies to night-time sampling at mountain-top sites; for this

transport model, these sites are sampled in model atmosphere layers well above the surface and, apparently, not as influenced by the vertical mixing closer to the surface. The value that these continuous, continental sites bring to the inversion is indisputable; for example, the Eastern Boreal North American region is not well constrained if we omit the Fraserdale site in favor of using only discrete sampling sites. It is also not recommended to mis-match the sampling protocols using all hours sampling from the model and mid-day hours from the observations for continuous sites. Sampling the model on a default basis would be a convenience for the modeler; however, choosing a default strategy for discrete sampling sites is problematic. Our results show that applying any default model sampling protocol at continuous or discrete sampling sites yields different solutions. Choosing mid-day hours from the continuous observations and matching model sampling hours with observations (as is done in the continental extension network inversion) appears to be the best solution with the transport model and observation sites used in this study.

4.3.2. Sensitivity to Classes of Observation Sites

Comparing inversions with and without the mountain-top sites in Europe, we find that differences are confined primarily to Europe and Boreal Asia. The lack of measurements in north central Europe and Boreal Asia make the shifting of the inversion results here not robust. This can also be seen in the posterior covariance structure in **Figure 2-5**, where the Europe – Boreal Asia covariance has a larger magnitude than any other regional pair outside of South America and Africa. It would take more observation sites in Boreal Asia and North Central Europe to resolve this.

Two aircraft observation time series are included in this inversion: the vertical profile at Briggsdale, Colorado and the high altitude western Pacific transit from Japan to Australia.

Removing the Colorado profile shifts the distribution of the North American carbon sink from Boreal to Temperate aggregated regions, but these shifts are within the posterior uncertainty of the control inversion using the continental extension (CE) network. Other North American vertical profiles were not used because they did not meet the completeness criteria for 2000-2004; it would be possible to include these vertical profiles with a more recent solution time span. More investigation is needed to determine if these profiles should be included in the inversion observation set, or instead used to validate the inversion solution as recommended elsewhere (e.g., Peters et al., 2007). The results of our experiments do not answer the outstanding question of whether atmospheric inversions using observations from surface sites alone can adequately characterize global carbon sources and sinks (Stephens et al., 2007). The use of the western Pacific transit observations is even more equivocal. We do not know if the transport is accurate at the ~10500 m altitude of these measurements. When we replaced the high altitude observations with the NOAA Pacific Ocean surface observations (not included in the base observation set because they did not meet completeness criteria), we see somewhat different results in the Pacific basin and surrounding land regions (larger sink in the North Pacific and larger sources in Tropical America, Boreal Asia, Tropical Asia and Australia) and reduced uncertainties. The resulting fluxes resemble the Pacific basin flux patterns found in other inversions (e.g., Baker et al., 2006; Peters et al., 2007). The only inversion variations for which Tropical Asia is a mean annual sink for 2001-2003 are those which include the high altitude aircraft observation data. An inversion variation using both the high altitude observations and the Pacific Ocean surface observations yields a result similar to that with only the surface observations, but with a neutral annual mean for Tropical Asia for 2001-2003. The surface ocean observations do provide more constraint in the form of reduced regional uncertainties in our study. The performance of the transport model at the altitude of the aircraft observations warrants further research. It is not possible, in the context

of this study, to determine whether these aircraft observations should be included in the inversion network.

4.4. Conclusions

This set of inversion experiments was intended to show the potential of using newly available North American well-calibrated carbon dioxide measurement time series in a global atmospheric inversion. We have explored the handling of these data sets in terms of matching selected hours of observations with the same hours of transport model output (co-sampling). In addition, we have challenged the experimental setup in terms of its sensitivity to inclusion of classes of observations. Following are specific recommendations for future network design and inversion experiment design:

For global atmospheric inversions using this transport model and transport fields, continental quasi-continuous observation measurements are best incorporated as sub-sampled for well-mixed atmospheric conditions (mid-day hours for surface sites and mid-night hours for elevated sites) with model sampling for matching hours. Developing monthly means of model output from all hours is not recommended when real observations of atmospheric carbon dioxide mixing ratios are used in place of a smoothed data product.

The issue of whether or not to include aircraft profile or high altitude aircraft observations in a global atmospheric inversion of surface observations is still not made clear from this work. The typical NOAA aircraft profiles, for example, sample at altitudes within the troposphere, where the transport modeling may be acceptable. When Pacific Ocean surface observations are substituted for the high altitude Western Pacific observations in this experiment, the flux solution is different, indicating a net annual sink in Tropical Asia. This differs from the

generally accepted view that when biomass burning emissions are included, as they are in the results reported here, Tropical Asia is a source. We cannot tell from our work which is correct. Evaluation of transport fields near the tropopause is warranted to determine if inclusion of the high altitude observations is a good idea.

The time span of this inversion experiment should be extended now with the availability of the recent additions to the network and the transport fields for more recent years. A longer-term results set would enable better comparison with other methods. A logical next step with this global atmospheric inversion method is the inclusion of a carbon dioxide column observation data product from satellite retrievals to improve the observation density in those parts of the world that are currently not well sampled.

Chapter 5

Virtual Tall Towers and Future Networks

5.1. Introduction

This chapter covers two aspects of the future potential of the North American carbon dioxide measurement network. The CE (continental extension) network introduced in Chapter 2 included four North American flux towers. In Chapter 4 we discussed the characterization of the modeled and observed mixing ratio records at continental surface sites and concluded that mid-day sampling was necessary in order to better match transport model capabilities with the observations. In this chapter the virtual tall tower (VTT) concept is introduced; this is a micrometeorological adjustment to continental surface layer carbon dioxide mixing ratio observations to approximate observations in the mid-day mid-continental boundary layer. If the transport model output is a better match to mid-boundary layer observations, then this adjustment can be used at flux tower sites to simulate tall towers at the same locations.

Also in this chapter a method for creating simulated observations is presented. These simulated observations are used to test future North American networks incorporating observation locations that either already were instrumented in 2009 or are plausible sites for future network expansion. Inversions with these future networks are used to test posterior flux uncertainty reduction provided by additional North American sites.

In section 5.2, the methods for producing the VTT micrometeorological adjustment and the simulated observations are described. Results are shown in section 5.3, and conclusions about future networks are given in section 5.4. The inversion variations for this Chapter are documented

in **Table 5-1**. The observation site-network cross reference can be found in the Appendix in **Table A-11**.

Table **5-1**: Inversion variations discussed in Chapter 5.

Inversion	Network	# of Sites	Description
Control Inversion	CE	91	Continental Extension Network inversion described in Chapter 2; model output co-sampled
Virtual Tall Tower Experiments			
No surface layer, continental continuous	V1	79	No North American surface layer continuous sites
VTT, No surface layer continuous	V2	86	VTT adjusted North American flux towers, highest level at LEF and WKT, and Tapajos flux tower added to V1 network
VTT, Add FSD	V3	87	VTT adjusted North American flux towers; FSD added to V2 network
VTT version of control inversion	V4	91	VTT adjusted North American flux towers; all other sites same as CE network
Simulated Network Inversions			
Future Network 1	F1	115	CE network plus 24 North American sites representative of sites active in 2009
Future Network 2	F2	123	F1 network plus 8 more North American sites

5.2. Methods

5.2.1. Virtual Tall Tower Micrometeorological Adjustment

The virtual tall tower (VTT) micrometeorological adjustment to the surface layer measured carbon dioxide concentration follows the mixed layer similarity theory for the vertical gradient of a scalar in the boundary layer (Wyngaard and Brost, 1984; Moeng and Wyngaard, 1989). This vertical gradient is characterized by the fluxes at the top and bottom of the boundary layer in terms of universal top-down and bottom-up gradient functions:

$$\frac{\partial C}{\partial z} = -g_b\left(\frac{z}{z_i}\right)\frac{\overline{wc_0}}{w_* z_i} - g_t\left(\frac{z}{z_i}\right)\frac{\overline{wc_{z_i}}}{w_* z_i} \tag{5-1}$$

where C is a conserved scalar, in this case the mixing ratio of carbon dioxide in the continental boundary layer, z is the height above the surface, g_b and g_t are the universal bottom-up and top-down gradient functions scaled by the boundary layer depth z_i, and w_* is the convective velocity scale given by $\left[\frac{g z_i}{\theta_0}\left(\overline{w\theta_0}\right)\right]^{\frac{1}{3}}$, where θ_0 is the mean virtual potential temperature, $\overline{w\theta_0}$ is the sensible heat flux at the surface and g is the gravitational acceleration. The turbulent surface and entrainment fluxes for the scalar are $\overline{wc_0}$ and $\overline{wc_{z_i}}$. We implement this at the flux tower locations using temperature, carbon dioxide mixing ratio, carbon dioxide and sensible heat fluxes, all measured at the tower, and a displacement height implied by the local vegetation. We then compute the difference between the mixing ratio measured at the tower height and the mixing ratio that would be observed at 400 m above the ground using

$$\Delta C = -\frac{\overline{wc_0}}{w_* z_i}\int_{z_m}^{z_{VTT}} g_b\left(\frac{z-d}{z_i}\right)dz - \alpha\frac{\overline{wc_0}}{w_* z_i}\int_{z_m}^{z_{VTT}} g_t\left(\frac{z-d}{z_i}\right)dz \tag{5-2}$$

where ΔC is the difference between the carbon dioxide mixing ratio measured at z_m, the measurement height (typically ~30 m), and the mixing ratio at z_{VTT}, the virtual tall tower height (400 m). d is the displacement height; g_b and g_t are the bottom-up and top-down gradient functions from the empirical fit of Wang et al. (2007). Boundary layer depth and the entrainment flux are not routinely measured; we calculate the boundary layer depth, z_i, as in Yi et al. (2001), and specify the fraction α of the turbulent carbon flux measured at the tower as an approximation of the entrainment flux at the top of the boundary layer. Our formulation is applied to mid-day hours only. Note that while there is no reason for the entrainment flux to be a fraction of the

surface flux, this is not unusual during the day, and in most convective boundary layers the entrainment flux contributes little to the mixing ratio difference between the surface layer and 400 m. We tested the adjustment for multiple years of observations at the WLEF tall tower in Wisconsin, where we can compare the difference in carbon dioxide mixing ratio between 30 m and 396 m to the adjustment from equation 5-2. We found it necessary to screen for minimal sensible heat flux, boundary layer depth, and convective velocity scale; in particular, we omit hours from the calculation where z_{VTT} (chosen to be 400 m) is close to the top of the boundary layer to minimize the contribution of the second term in equation 5-2 containing the unknown entrainment flux. Overall we can employ the calculation for about half of the 12-16 LST hours annually using this screening due to a combination of missing observation data and failure to meet screening criteria. At the WLEF tall tower, we calculated the adjustments to the 30 m observed mixing ratio to approximate the top of the tall tower and compared the results with the 396 m observed mixing ratio. Over a period of 6 years in the test, the calculated hourly adjustment in summer was ~1 ppm with a bias of ~ -0.2 ppm; in spring and fall the adjustment was less than 0.5 ppm with a bias of less than 0.1 ppm; in winter, the mean hourly adjustment is close to zero with a bias of ~0.5 ppm and a higher rate of failure to meet the minimum criteria for calculation.

5.2.2. Simulating Monthly Carbon Dioxide Mixing Ratios

We also test the uncertainty reduction that would be obtained by including in the observation network an additional array of North American observation sites. The sites chosen are primarily those that have well-calibrated carbon dioxide mixing ratio measurements in 2009 (simulated sites in **Figure 5-1** and **Table A-6**).

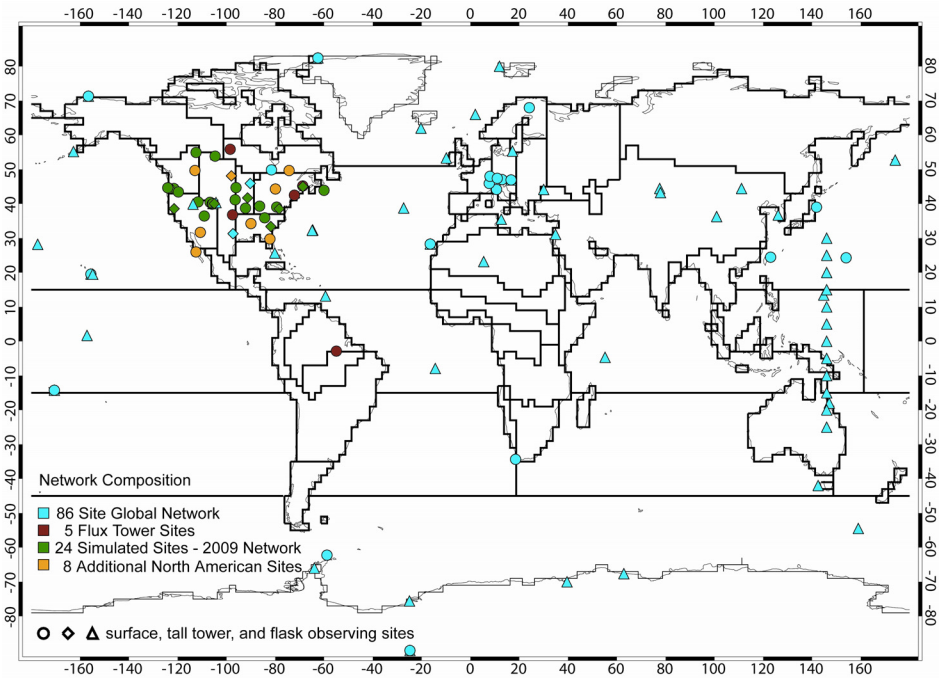

Figure **5-1**: Symbols mark observing sites in the future network tests. Symbol color indicates the network: 91-site continental extension (cyan); 24 sites added in 115-site future network (green); 8 additional sites in the 123-site future network (gold). Symbol shape indicates the type of observation: quasi-continuous (circle), tall tower (diamond) and discrete (triangle).

We simulate monthly mean observations for these additional sites treating them as quasi-continuous surface layer or tall tower sites. The intent is to simulate a seasonal cycle with amplitude and trend consistent with nearby observations; the method should not be expected to predict actual local observations. Monthly mean mid-day observations are created by adjusting the monthly mean model output of the composite background fluxes (terrestrial, ocean , fossil, and biomass burning) at these locations, using the default selection protocol (12-16 LST every day). We calculate a 6-parameter fit of the modeled monthly mean mixing ratios using the IDL curvefit routine (ITT Visual Information Systems) with a user-supplied function as

$$C_{\mathrm{mod}} = a_0 + a_1 t + a_2 \cos(2\pi t) + a_3 \sin(2\pi t) + a_4 \cos(4\pi t) + a_5 \sin(4\pi t) \qquad (5\text{-}3)$$

where C_{mod} is the monthly mean time series of mid-day model samples of background fluxes for five years at the site being simulated and t is a time variable in years (0 to 5). We substitute the offset and trend parameters (a_0, a_1) for the GLOBALVIEW (GLOBALVIEW-CO2, 2007) marine boundary layer (MBL) fit interpolated to the same latitude as the observing site being simulated. Finally, we adjust the amplitude of the harmonics to account for the amplitude of the background SiB3 terrestrial flux being larger than observed in North America (Law et al., 2008). The amplitude adjustment is based on comparison to nearby observation sites for which we do have existing data. The simulated observation time series is then

$$C_{sim} = a_{0_{mbl}} + a_{1_{mbl}} t + \beta\left[a_2 \cos(2\pi t) + a_3 \sin(2\pi t) + a_4 \cos(4\pi t) + a_5 \sin(4\pi t)\right] \quad (5\text{-}4)$$

where β is the multiplier adjustment for the harmonics. Monthly uncertainties are derived from the monthly variability of the monthly mean model samples of the composite background fluxes, and are at least as large as the monthly variability of actual observations at locations of similar type. The offset, trend and harmonics multipliers used for each simulated site are listed in **Table 5-2**. For the three sites with offset and trend marked "n/a" in **Table 5-2**, some observations exist for the 2000-2005 time period of our experiment. These incomplete time series were gap-filled with the GLOBALVIEW-CO2 (2007) data product. The amplitude multiplier adjustment for these sites is the computed ratio of the harmonics of the gap-filled and composite background model output fits. The simulated observations will be smoother than the actual observations and will not take into account short-term deviations from the seasonal cycle. The simulated time series (not shown) are better matched to observations at the northern forested sites which have distinct seasonal cycles.

Table **5-2**: Parameters for Simulated Observation Time Series. Site No. is a reference to the sites in Table A-6. The simulation parameters are Offset (ppm), Trend (ppm yr^{-1}), and Amplitude Multiplier (β). See text for description of the simulation algorithm and n/a entries.

	Location		Simulation Parameters		
Site No.	Lat.	Lon.	Offset	Trend	β
1	54.95	-112.47	368.56	2.09	0.55
2	53.87	-104.65	n/a	n/a	0.55
3	45.03	-68.68	n/a	n/a	0.56
4	44.72	-96.09	368.11	2.23	0.55
5	44.67	-124.07	368.11	2.23	0.55
6	44.50	-123.55	368.10	2.23	0.55
7	44.45	-121.56	368.09	2.23	0.55
8	43.45	-119.72	368.08	2.24	0.55
9	43.93	-60.02	n/a	n/a	0.63
10	41.72	-91.35	368.21	2.22	0.65
11	40.56	-111.64	368.34	2.19	0.85
12	40.45	-106.73	368.36	2.19	0.85
13	40.05	-105.58	368.42	2.17	0.85
14	40.05	-105.01	368.42	2.17	0.85
15	41.48	-96.44	368.27	2.21	0.55
16	39.91	-105.88	368.44	2.17	0.85
17	39.32	-86.41	368.51	2.15	0.55
18	39.06	-79.42	368.54	2.14	0.55
19	38.74	-92.20	368.57	2.13	0.55
20	38.57	-121.47	368.59	2.13	0.55
21	38.53	-78.43	368.59	2.13	0.80
22	36.46	-109.10	368.72	2.09	0.85
23	35.93	-84.33	368.73	2.09	0.55
24	33.40	-81.83	368.79	2.06	0.55
25	49.71	-112.93	368.44	2.12	0.55
26	49.69	-74.34	368.44	2.12	0.80
27	48.13	-97.98	368.37	2.14	0.80
28	44.32	-79.93	368.09	2.24	0.80
29	34.25	-89.97	368.77	2.07	0.60
30	31.70	-110.72	368.82	2.04	0.60
31	29.77	-82.20	368.82	2.04	0.60
32	26.00	-112.50	368.74	2.08	0.60

5.3 Results

5.3.1. Testing the Virtual Tall Tower Concept

For the four North American flux towers in the 91-site CE (continental extension) network, we calculate the adjustment to approximate the 400 m carbon dioxide mixing ratio for each mid-day hour with available data meeting minimum criteria. Monthly means of the VTT adjusted observations are calculated and shown in the first column of **Figure 5-2** for three North American flux towers (rows A, B and C) and the observed mixing ratio of carbon dioxide at the 396 m level of the WLEF tall tower (row D). **Figure 5-2** also shows the differences in monthly means calculated from the VTT adjusted mixing ratios compared to the surface values (column 2), and the difference in model samples from 400 m (between layers 2 and 3 in the model atmosphere) and the surface (column 3). In row D of **Figure 5-2**, the third column shows the differences between model samples at 396m and 30m at the WLEF tall tower to show this difference at an existing tall tower.

The vertical difference in the monthly mean mixing ratios of observations and model samples is similar for the tall tower. The difference between the VTT adjusted observations and the model samples at the flux towers is of similar order of magnitude. It cannot be determined from this simple comparison whether dissimilarities in the seasonal pattern and amplitude of the differences are due to failure of the VTT adjustment algorithm or possible problems with the vertical gradient in the model transport.

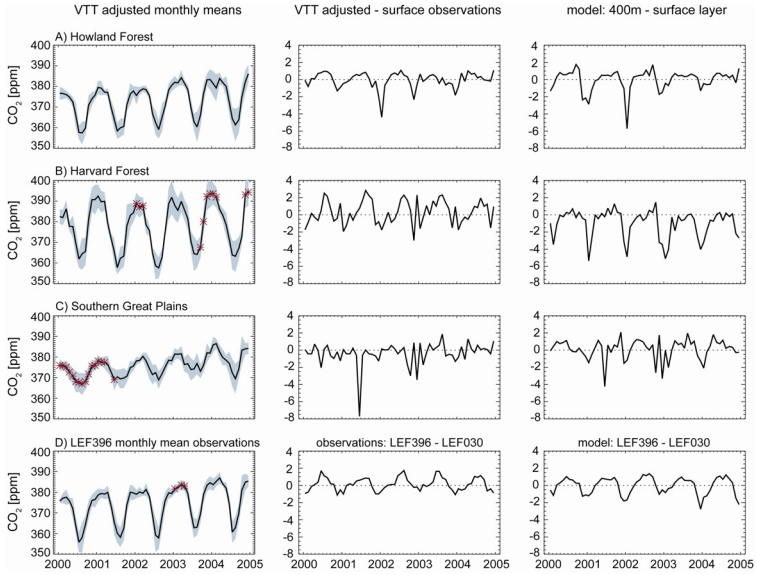

Figure **5-2**: The first column shows example monthly mean VTT adjusted mixing ratios at flux towers: rows A Howland Forest, B Harvard Forest, and C Southern Great Plains. The second column shows the difference between monthly mean VTT adjusted and surface layer mixing ratios. The third column shows differences in modeled mixing ratios at 400 m and surface layer at the same locations. Row D shows the equivalent differences in observations and model samples between the 396 m and 30 m heights of the WLEF tall tower. All units are ppm CO_2. Data points marked with asterisks (*) indicate gap-filled data.

We test the virtual tall tower concept with inversions for four networks (**Table 5-1**), varying the number of sites in North America: the 79-site V1 network with no continuous observing sites in North America, the 86-site V2 network adding the observations at the top of the Wisconsin (LEF) and Texas (WKT) tall towers and the VTT adjusted observations at the four North American flux towers, the 87-site V3 network adding Fraserdale (FSD), and the 91-site V4 network that has the same sites as the continental extension (CE) network, but with VTT adjusted time series for the four flux towers. See also the network-site cross reference in **Table A-11**. The results are shown in **Figure 5-3**, where the thick black line and shading represent the results of the control inversion using the 91-site continental extension network (CE) from Chapter 2.

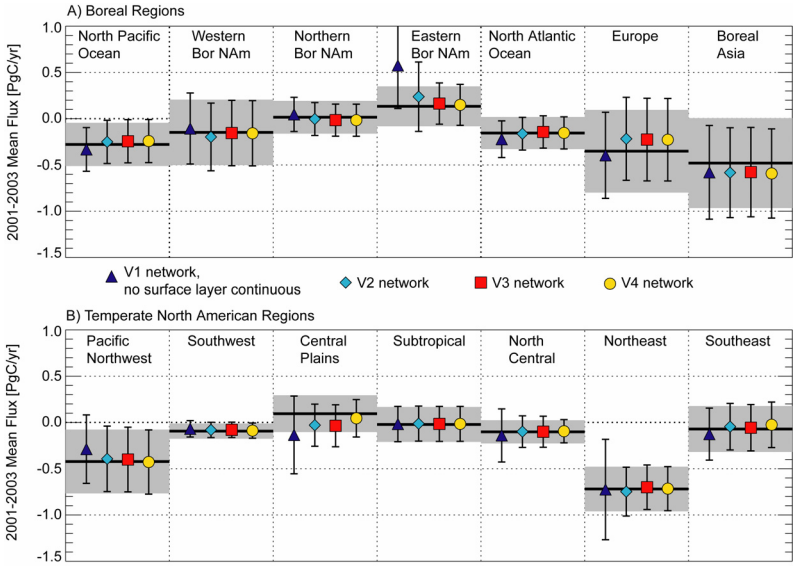

Figure **5-3**: Mean annual fluxes and uncertainties for 2001-2003 for A) Boreal regions and B) Temperate North American regions illustrating the use of surface layer and virtual tall tower adjusted carbon dioxide mixing ratio observations. Solid line and shading represent the 91-site CE network inversion. Symbols indicate the network used in the inversion: 79 sites (V1, triangles), 86 sites (V2, diamonds), 87 sites (V3, squares) and 91 sites (V4, circles). All units are Pg C yr^{-1}.

The V1 network inversion result (triangles in **Figure 5-3**) demonstrates the increase in uncertainties in the regions where the continental, continuous North American sites are located in the CE network (for example, different flux solutions for Eastern Boreal and Central Plains regions and increased uncertainty in many of the Temperate North American regions). Adding the VTT-adjusted flux tower sites and the top level of the LEF and WKT tall towers in the V2 network (diamonds in **Figure 5-3**) brings the Eastern Boreal and Central Plains results closer to the control inversion and significantly reduces the uncertainty of Eastern Boreal, Central Plains, North Central and Northeast regions where these introduced observing sites are located. The Fraserdale site included in the V3 network is seen to be essential for constraining the flux in the

Eastern Boreal region (squares) and the two lower levels of the tall towers used (30 m and 122 m) help to reduce the uncertainty in the Central Plains and North Central regions. Overall the VTT version of the CE network (V4, circles) results in a solution very close to the fluxes and uncertainties of the control inversion, with some redistribution of flux in Europe, Boreal Asia and the North Pacific and North Atlantic Oceans.

5.3.2. Inversions for Future Networks

We conduct a further experiment with simulations of monthly carbon dioxide mixing ratio time series at sites that have well-calibrated carbon dioxide measurements in 2009 to see what affect additional sites could have on our inversion. These are flux towers, tall towers, and instrumented sites located in complex terrain. We seek to determine the degree to which this expanded network will resolve the source/sink pattern within North America and reduce the uncertainty of the inversion solution, while not affecting the global distribution of sources and sinks in any significant way. We simulate monthly carbon dioxide time series for 2000-2004 at these new observation locations as well as at a few additional potential quasi-continuous observation sites in order to test the ability of these data to further reduce the uncertainty of the flux solution for the sub-regions in North America using this inversion method. There are two inversions in this part of the experiment (**Table 5-1**): the 115-site F1 network with the additional 24 sites in operation in 2009, and the 123-site F2 network with 8 additional sites at locations chosen to improve observation coverage. These 8 additional sites are at locations which are planned or have had observation programs in the past, although not necessarily well-calibrated carbon dioxide measurements.

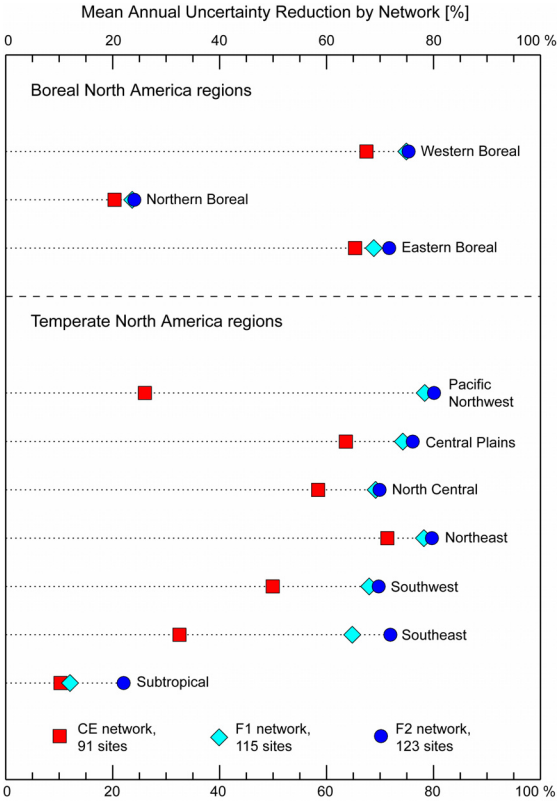

Figure **5-4**: Mean annual uncertainty reduction (in percent) for 2001-2003 for North American regions in three networks: 91-site CE network (squares); 115-site F1 network with simulated observations from 24 North American sites (diamonds); 123-site F2 network with simulated observations from 8 additional North American sites (circles).

Figure 5-4 shows the 2001-2003 mean uncertainty reductions (defined in Chapter 2 as $1-\frac{\sigma_{post}}{\sigma_{prior}}$, expressed as percent) achieved relative to the *a priori* uncertainty specifications.

Addition of observing sites in the F1 network in the Pacific Northwest region results in a significant reduction in uncertainty in the Pacific Northwest and a lesser reduction in the Western Boreal region. The uncertainty reduction due to additional observations is also marked in the Southeast. The uncertainty of the Subtropical region is not improved until a site is added there in the 123-site F2 network inversion. More than two sites in a region does not reduce the uncertainty much, but the degree to which these data would truly be redundant has yet to be tested with true

observations. This would be a good test of the inversion methodology for a more recent time period when there are more observations available.

Figure 5-5 compares the 47-region posterior annual covariance matrices for 2002 for the control inversion (CE network) and the F1 network. The region numbers on these charts are referenced to **Figure A-1** and **Table A-1**. Boreal North America is composed of regions 5-7 and North America, regions 8-14. Within North America variances are reduced in many regions: Pacific Northwest from 0.11 to 0.01, Southeast from 0.06 to 0.02, Northeast from 0.05 to 0.03, and the Central Plains from 0.04 to 0.02 $(\text{Pg C yr}^{-1})^2$. Changes in variances outside of North America are not significant (largest change is -0.03 $(\text{Pg C yr}^{-1})^2$ in Northern Temperate South America). A few significant covariances remain among the North American regions compared to the covariance matrix from the control inversion. Only the covariances among Boreal regions and between the Eastern Boreal and Northeast regions have magnitude larger than 0.01 $(\text{Pg C yr}^{-1})^2$ and none of these is of magnitude greater than 0.02 $(\text{Pg C yr}^{-1})^2$. Covariances between the North American regions and the remainder of the global regions all have magnitude smaller than 0.01 $(\text{Pg C yr}^{-1})^2$ except for the covariances between Western Boreal-Boreal Asia (-0.042 $(\text{Pg C yr}^{-1})^2$), Western Boreal-Temperate Asia (-0.015 $(\text{Pg C yr}^{-1})^2$), Eastern Boreal-Boreal Asia (+0.010 $(\text{Pg C yr}^{-1})^2$), and Subtropical-Central America (-0.012 $(\text{Pg C yr}^{-1})^2$). The 2002 posterior annual covariance matrix for the 123-site F2 network inversion is very similar to the F1 network inversion result shown at the bottom of **Figure 5-5**.

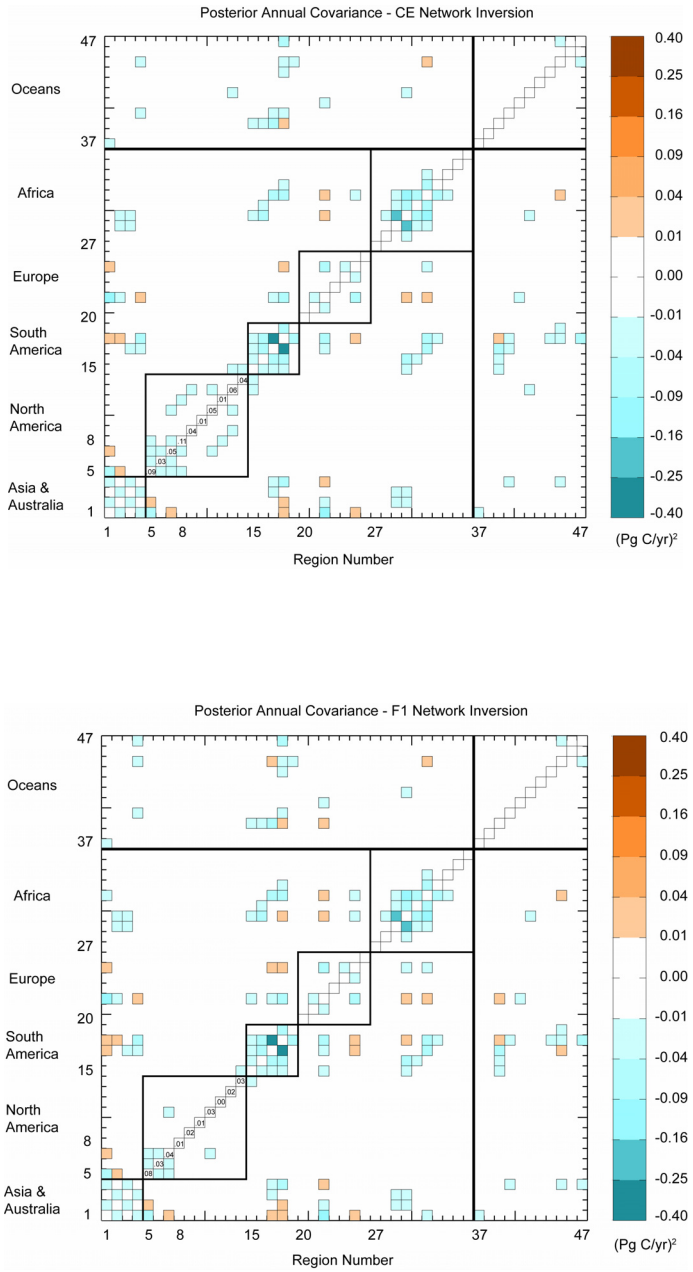

Figure **5-5**: Posterior annual covariance matrix for 2002 for CE network inversion (top) and F1 network inversion (bottom). Region number references are in Figure A-1 and Table A-1. Variances in $(\text{Pg C yr}^{-1})^2$ are shown on the diagonal for North American regions. Shaded boxes contribute ≥ 0.1 Pg C yr^{-1} to the region uncertainty.

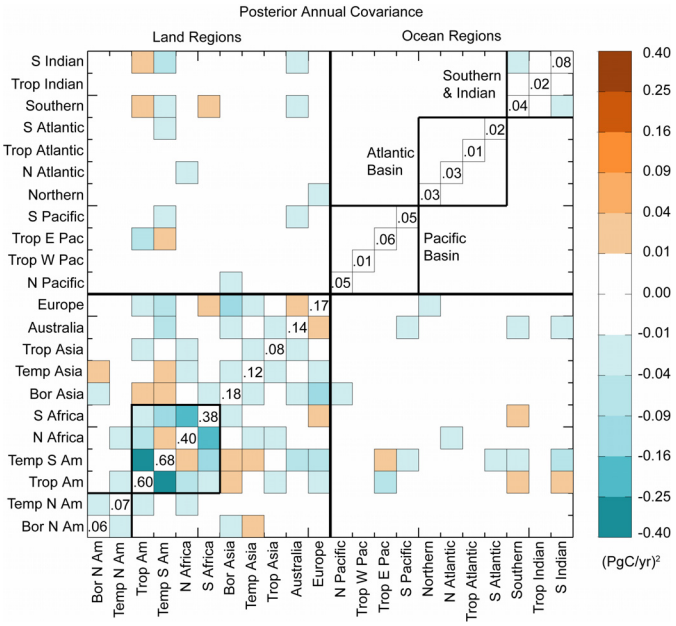

Figure **5-6**: Posterior annual covariance matrix for 2002 for the 115-site F1 network for 22 aggregated land and ocean basin regions. Variances in $(\text{Pg C yr}^{-1})^2$ are shown on the diagonal. Shaded boxes contribute ≥ 0.1 Pg C yr^{-1} to the aggregated region posterior uncertainty.

The posterior covariance matrix for 2002 annual fluxes for the aggregated TransCom regions (**Table A-1**) and the 115-site F1 network is presented in **Figure 5-6**. Variances are reported on the diagonal. Off-diagonal values are annual covariances indicating the extent to which the inversion determines the adjustments to the annual fluxes for each region as independent of the other regions. The shading of off-diagonal values indicates covariances which impact the annual uncertainty of these aggregated regions by 0.1 Pg C yr^{-1} or more. The ocean regions are for the most part independent by this measure, as is also true with the CE network inversion (see **Figure 2-5** in Chapter 2). The variance for Boreal North America is reduced from

0.09 to 0.06 (Pg C yr^{-1})^2 relative to the CE network inversion; the Temperate North America variance is reduced similarly from 0.15 to 0.07 (Pg C yr^{-1})^2 ; the covariance between them is changed from -0.05 to -0.02 (Pg C yr^{-1})^2 (area bounded by heavy black line in lower left corner of **Figure 5-7**). With the exception of small dependences between Boreal North America and Boreal and Temperate Asia, and between Temperate North America and Tropical America and North Africa, North America is independently constrained with this network. This network does not, however, constrain the South American and African regions (bounded by heavy black lines in **Figure 5-7**); the largest covariances are there: -0.39 (Pg C yr^{-1})^2 between Tropical America and Temperate South America, -0.21 (Pg C yr^{-1})^2 between Northern Africa and Southern Africa, -0.12 (Pg C yr^{-1})^2 between Temperate South America and Southern Africa. The next largest covariance is -0.10 (Pg C yr^{-1})^2 between Boreal Asia and Europe. As with the CE network inversion, we cannot justify the inversion solution for sub-regions within South America and Africa with the networks tested. Adding the 8 sites in the 123-site F2 observation network does not significantly change the aggregated posterior covariance or the regional posterior covariance.

5.4. Discussion

5.4.1. Use of the VTT Adjustment at Flux Tower Sites

This adjustment, which can be calculated at flux tower sites with standard observation data, is intended to approximate mid-boundary layer carbon dioxide mixing ratios from surface values. Assuming that the transport above the model surface layer is more accurately modeled and that the calculated mid-boundary layer value is representative of a larger area than the immediate vicinity of the flux tower, this adjustment could be used to simulate tall towers at flux tower sites. The VTT adjusted monthly mean carbon dioxide concentrations at the flux tower

sites we use here are within 1-2 ppm of surface values for most months of the year; model samples for the same hours with this transport model show differences on the order of 1 ppm in the summer and 2 ppm in winter. This difference in model and adjusted surface mixing ratios appears not to cause large excursions in the inversion results. The VTT adjustment can only be calculated during mid-day hours with convective conditions, so selected hours time-of-day sampling must be used in both observations and model. The availability of the measured quantities required to calculate the VTT adjustment and the screening for appropriate convective conditions reduces the number of hours of observations by about one half, but this also appears not to have a large effect on the inversion results. Consideration must, however, be given to the bias in favor of convective conditions and an increased uncertainty in the adjusted observations (not implemented here) due to the uncertainties of the observed quantities used in the adjustment and the approximations used. Depending on the boundary layer representations in other transport models, this may be a viable approach if middle levels of the boundary layer are true to the observations even when the surface representation is not.

5.4.2. Future Networks

We show here that the 2009 network of 115 observation sites does reduce the uncertainty in the regional inversion solution in North America, with the exception of the Subtropical region, which has no observation sites in this network. Using either half of the additional 24 observation sites in the F1 network yields marginally different flux results but very little difference in uncertainty (not shown). The difference in the solutions appears to be the extent to which regions within North America are constrained by the local observations. A few observation sites (2-3) appear to be adequate to constrain regions of the size used in this experiment, but this should be tested with the real observation data when it is available. For this experiment, there are adequate

observations to constrain Temperate North America. The increased density of observations available in the F1 network makes regional atmospheric inversions with finer spatial resolution more feasible.

5.5. Conclusions

This set of inversion experiments was intended to show the potential of using newly available North American well-calibrated carbon dioxide measurement time series in a global atmospheric inversion. The virtual tall tower concept was also implemented to study how mid-boundary layer continental measurements might be more compatible with transport model samples. Following are specific recommendations for future network design and inversion experiment design:

The virtual tall tower adjustment to surface observations, which can be calculated using atmospheric and carbon flux variables at flux tower sites, may be appropriate for use with transport models that represent the continental boundary layer ~400 m above the surface better than at the surface. The adjustment brings an added uncertainty from the observed variables used in the algorithm and the limited conditions under which the adjustment can be calculated. Inversion results using the VTT adjusted observations are comparable to those using mid-day surface layer observations at the same locations in our experiment.

The addition of multiple sites in North America with well-calibrated, quasi-continuous carbon dioxide measurements offers opportunities for regional atmospheric inversions at finer spatial resolution as well as the global atmospheric inversion studied in this paper. Having a few observations in each region of the size used here appears to be enough to resolve the flux. Using half of the new observation sites gives uncertainty reductions nearly equivalent to the whole set.

The time span of this inversion experiment should be extended now with the availability of the recent additions to the network and the transport fields for more recent years. A longer analysis period of inversion results would enable better comparison with other methods, and might permit inclusion of some of the sites simulated in the future networks in this work. A logical next step with this global atmospheric inversion method is the inclusion of a carbon dioxide column observation data product from satellite retrievals to improve the observation density in those parts of the world that are currently not well-sampled.

Chapter 6

Summary and Future Work

In this work we have used a Bayesian synthesis global atmospheric inversion to infer carbon sources and sinks for sub-continental regions and ocean basins for the period 2001-2003. The experiment includes several variations of the inversion intended to explore both methods and network design. We find that including continental measurements of carbon dioxide can improve the inversion results by reducing the uncertainty of the spatial and temporal distribution of continental carbon sources and sinks. Including continental sites in such an inversion requires simple sub-sampling of well-calibrated carbon dioxide time series for well-mixed conditions, but otherwise these data appear no more challenging to use than the "remote" data traditionally used in atmospheric inversions.

Section 6.1 covers methodological improvements introduced in this work and remaining issues; section 6.2 highlights findings about network design. Recommendations for future work are in section 6.3. The inversion results from the continental extension network in Chapter 2 have been contributed to the North American Carbon Program Interim Synthesis activity and to a 2009 comparison of contemporary inversions by the Atmospheric Transport Model Intercomparison (TransCom) working group.

6.1. Methodology

As with any difficult problem, initial solution methods often involve simplifications and assumptions that make the problem more tractable. The initial approaches are avenues to learning about the problem and the significance and consequences of the assumptions in the methods.

Over time the methodological issues are addressed or new methods are applied. We follow in a long line of researchers using the least squares analytic linear solution method, knowing that criteria for use of the method are not met. We hope, as Tarantola (2005) warns, that we have not violated the criteria too much.

We have addressed some of the simplifications in the method that have been used in previous work. Instead of re-using a single year of meteorological driver data in a climatology approach to atmospheric transport, we have used transport fields appropriate for the years of the analysis. Using annually varying atmospheric transport required modification of the inversion to accommodate annual variability in all transport response functions. In addition, we used carbon dioxide time series data directly in computing monthly means at each site and co-sampled the transport model output to match the observation hours. Earlier practices of using a smoothed data product and reusing transport fields may have been appropriate when the intent of the inversion was a mean annual or mean seasonal flux solution.

We accepted the background fluxes as fact; any adjustments calculated in the inversion are assumed to be adjustments to the background terrestrial and ocean surface fluxes. Although we have implemented seasonal variation in the fossil fuel emissions and tested the sensitivity of our results to non-varying vs. seasonally varying emissions, the underlying spatial distribution we used is out-of-date for the inversion analysis time period. Errors in this distribution are aliased in the adjustments to the terrestrial background flux. Our results were not sensitive to the choice of terrestrial background flux in regions of high observation density. The SiB3 and CASA fluxes do not agree in some regions on amplitude or timing of seasonal cycles, but this inversion method could correct both to a common solution in regions that are well-represented in the observation network. This is consistent with an overall finding that regions with adequate representation in the global observation network are generally not overly reliant on either the terrestrial background flux used or the magnitude of the prior flux uncertainties.

Some aspects of the method have not been addressed here; the principal assumptions that the errors in the data, forward model, and priors are Gaussian and not correlated are still outstanding. We have used diagonal prior covariance matrices for both carbon fluxes and atmospheric mixing ratio data, knowing that this cannot be completely correct. If this assumption were strictly true, then the posterior flux covariance matrix and the data residuals should also describe Gaussian distributions. The data residual distribution is more peaked than Gaussian and centered at a few tenths of a ppm below zero, and the posterior flux covariance still holds off-diagonal elements, primarily between regions not well represented in the observation network. This would suggest that a general linear solution method would be more appropriate, using prior covariance matrices for both flux and data that include the appropriate covariances in space and time. We have chosen to follow past practitioners in following a solution technique that is not quite adequate, but useful for its relatively straightforward application and informative solution products. The main consequence of using the method is that the posterior flux uncertainties we report are dependent on our assumed uncertainties in the prior fluxes and data. Neglecting the prior covariances may cause our posterior uncertainties to be too small. Additional research with improved methods is required to address this.

We also assume that the underlying pattern of flux distribution within each region is correct, so that the flux solution applies to a whole region. This is also unlikely to be true, but this assumption should become less important as the spatial resolution of the solution is pushed toward smaller regions where the biota are likely to respond in similar ways to weather and climate patterns. In this work, we show that sub-continental regions can work in North America with its relatively dense observation network; elsewhere the global network is too sparse to support this sub-continental spatial resolution.

6.2. Network Design

We have successfully incorporated carbon dioxide time series, calibrated to global standards, from flux towers, in the observation network used for global atmospheric inversions. We sub-sampled the observations from these sites for mid-day hours to choose conditions when the continental boundary layer is most likely to be well-mixed, conditions most likely to be simulated best by the transport model. An inversion for the network including four flux towers in North America resulted in a reduction of the 2001-2003 mean annual North American sink by 15% to about 0.4 PgC yr^{-1} compared to an inversion for a typical global network of quasi-continuous and discrete sampling sites.

We recommend co-sampling the model atmosphere at the same times as the observations when using actual observations rather than a smoothed data product. Inversion variations using transport model samples from all hours, and even hours selected using default sampling criteria, when matched against our observation data set, produced anomalous inversion results, especially for regions with continental quasi-continuous observing sites.

We used an empirical approach to the question of balancing region size and number of observation sites. We defined ten regions in North America, for example, based roughly on climate, ecosystem factors, and the distribution of available observations. We did not design the network, but chose existing sites or those that have the potential of being instrumented for carbon dioxide measurement and calibration to global standards. Our results showed that the incremental uncertainty reduction in source/sink estimation is local to the region in which additional sites are located. In the previously unrepresented Northeast region, adding two sites resulted in reduced uncertainty for the region and reduced posterior covariance with other regions within North America. The two sites are in different forest ecosystems, with carbon dioxide time series that

differ in amplitude of seasonal cycle and five-year growth trend; together they provide complementary information to the inversion solution.

On the other hand, including co-located surface sites provides little new information to the inversion; there will appear to be a reduction in posterior uncertainty but little modification of the flux. In the interests of increasing the independence of the observations and their errors, we recommend that locations not be duplicated in the observation network unless it is clear that the observation sites sample different air masses, such as those separated vertically by a thousand meters or more. Much finer transport model resolution than we used would be required to distinguish surface sites near each other. For example, in our inversions, the two discrete sampling sites in Bermuda cannot both be modeled well even when their shared transport model grid cell is co-sampled to match the different sampling times at these two sites.

We also tested classes of sites other than continental flux towers that have been challenging to past inversions, and do not have enough information to draw firm conclusions about the results. We included one North American aircraft profile and one high altitude western Pacific Ocean transit time series in our global observation network. The North American profile increases the North American sink estimate slightly and reduces the posterior uncertainty by a very small amount. There are more North American aircraft profile sampling observations now that could be included in future testing. They should be useful either as observation sites in the inversion network or as validation of the inversion output. The high altitude Pacific Ocean transit series, however, provides a different solution for the fluxes in the Pacific basin regions compared to a network using surface observations from ships in the central Pacific. Using both the high altitude and surface data from the Pacific Ocean yields a result that mostly closely resembles the inversion variation using the surface data and not the high altitude data. Further validation of the transport model performance at high altitudes is recommended.

Including or omitting the mountain top sites in Europe causes redistribution of posterior flux between north central Europe and Boreal Asia. With no observation sites in Boreal Asia in the networks tested, there is not enough information for the inversion to separate the flux contributions of north central Europe and Boreal Asia consistently. The balance of estimated fluxes between these two areas shifts with different network configurations. Defining sub-regions within Boreal Asia might help to localize the least constrained areas; this would be helpful for network design and the placement of future observation sites, but would not necessarily improve the estimated flux solution with the networks used here. Including north central Europe in the Boreal Asia region, rather than in the aggregated Europe region might also result in more consistent results with the observation networks tested in this experiment.

The scarcity of measurement sites in Southern Hemisphere land regions, especially South America, makes it unlikely that we can distinguish between South America and Africa in our inversion solutions. The one South American site we used is not modeled well by the terrestrial background flux. Improvements are needed in the terrestrial flux products, especially in areas with sparse measurements. More continental observation data are needed in the Southern Hemisphere in general. With the current networks, the regions with low observation density are where the background fluxes and tighter prior uncertainties are necessary to constrain the inversion solution to be biogeochemically realistic.

A final note about networks concerns the requirement of our method for gap-free observation time series; gaps will occur in the actual measurements due to instrument problems and changes in measurement programs. Although gap-filling strategies can be used to estimate the missing observations, this inversion method requires long-running observation sites. The alternative of using shorter windows of time for inversions with changing network components makes determination of long-term trends more difficult.

6.3. Future Directions

This inversion as currently configured should be run for a longer analysis period to determine long term trends and annual variability. This would enable better comparisons with other contemporary inversions. A more recent analysis period would also permit the use of many of the continental sites now operational that we tested in our future network experiments. The use of the aircraft profile observations could also be more thoroughly tested. It may be possible to reduce the region sizes in North America with a denser observation network. The virtual tall tower adjustment could also be investigated at more sites and with more atmospheric transport models to develop more general criteria for use. Improved versions of the background fluxes and finer spatial resolution of the transport model can lead to mutual improvement of the terrestrial and ocean fluxes and the transport models. Continued comparisons of results using alternative background fluxes and different transport models as in the continuing TransCom model comparison projects is necessary as we are not yet at the stage where the results from different methods converge.

The method we used could be used to conduct a network optimization exercise to choose surface observation sites for future networks. There have been other network optimization efforts (e.g., Gloor et al., 2000; Patra et al., 2003; Patra et al., 2006; Rayner, 2004), but these do not include current observation sites nor do they deal with the temporal or spatial solution structure we have used here. Finding an appropriate balance between the density of the observation network and the size of regions for flux estimation is still a goal.

This inversion should also be modified to make use of new types of observations, especially column CO_2 data products from the Total Column Carbon Observing Network (TCCON) (Washenfelder et al., 2006; Yang et al., 2007) and the satellite observations from programs designed for CO_2 retrieval: Greenhouse gases Observing SATellite (GOSAT, also

know as IBUKI) (Yokota et al., 2004) and a possible replacement of the failed Orbiting Carbon Observatory (OCO) (Crisp et al., 2004). Advantages of incorporating column CO_2 into the observing network include the increased global coverage not likely with a surface network and resolution of the question of whether surface sources/sinks can be adequately characterized with only surface observations (Stephens et al., 2007). Experiments with other inversion methods are already being conducted with simulated column CO_2 products. The inversion method we used, while violating some criteria for applicability, is expedient for use as a first approach, and belongs in the inversion toolkit for initial assessments of new observation types and transport models, and perhaps also to supply boundary conditions, prior fluxes or covariances to alternative methods.

There is much work yet to do before we can be confident in our diagnosis of global carbon sources and sinks at a resolution that will be useful for either variability attribution or regulatory monitoring.

Appendix

Common Reference Material

A.1. Regions and Aggregated Regions

The spatial resolution of the inversion is shown in **Figure A-1**. In **Table A-1**, the land regions are named and aggregated regions that conform to the Atmospheric Transport Model Intercomparison Project (TransCom) regions are defined.

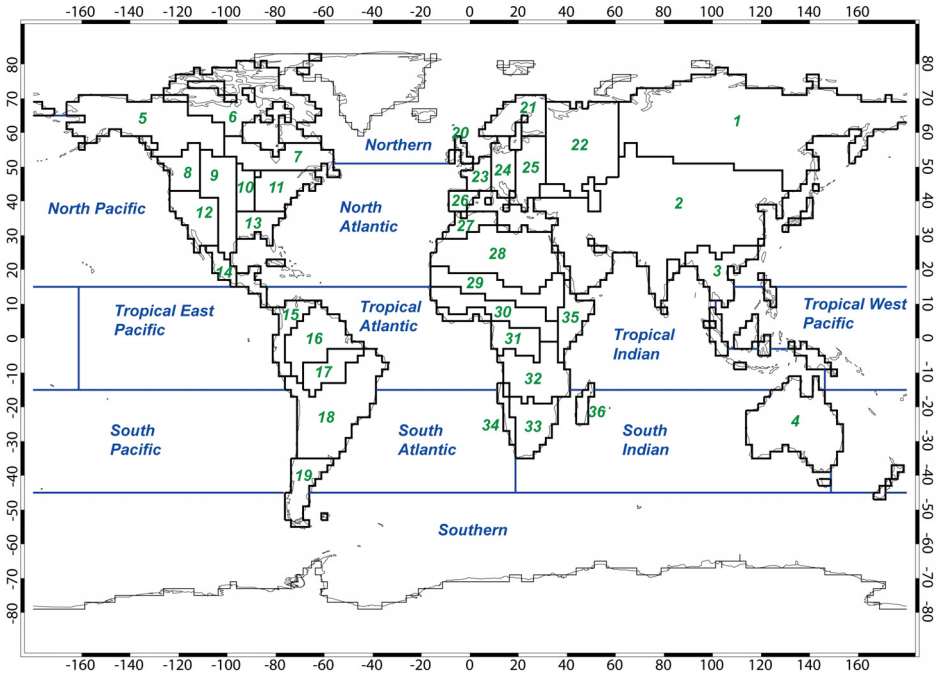

Figure **A-1**: Map depicting the 36 land regions and 11 ocean regions. Numbers in the land regions correspond to region names in Table A-1.

Table **A-1**: Region Names and Definitions of Aggregated Regions. The aggregated land regions correspond to land regions used in the Atmospheric Tracer Transport Model Intercomparison Project (TransCom); ocean regions are consistent with the TransCom ocean regions.

Aggregated Region	Map Number	Region Name
Boreal North America		
	5	Western Boreal North America
	6	Northern Boreal North America
	7	Eastern Boreal North America
Temperate North America		
	8	Pacific Northwest North America
	9	Central Plains North America
	10	North Central North America
	11	Northeast North America
	12	Southwest North America
	13	Southeast North America
	14	Subtropical America
Tropical America		
	15	Central America
	16	Northern Amazonia
	17	Southern Amazonia
Temperate South America		
	18	Northern Temperate South America
	19	Southern Temperate South America
Northern Africa		
	27	Mediterranean Coast Africa
	28	Northern Arid Africa
	29	Northern Dry Savanna
	30	Northern Mesic Savanna
	35	Horn of Africa
Southern Africa		
	31	Western Forest Africa
	32	Southern Mesic Savanna
	33	Southern Dry Savanna
	34	Southern Arid Africa
	36	Madagascar
Boreal Asia	1	Boreal Asia
Temperate Asia	2	Temperate Asia
Tropical Asia	3	Tropical Asia
Australia	4	Australia
Europe		
	20	British Isles
	21	Scandinavia
	22	North Central Europe
	23	Western Europe
	24	Interior Central Europe
	25	Eastern Europe
	26	Iberia

A.2. Observation Data Sources

Table **A-2**: Key to Observation Measurement Program Agencies

Acronym	Agency
AEMET	Meteorological State Agency of Spain
CESI RICERCA	Environment and Sustainable Development Department (Italy)
CNR-ICES	International Center for Earth Sciences (Italy)
CSIRO	CSIRO Marine and Atmospheric Research GASLAB (Australia)
DOE LBNL	Department of Energy Lawrence Berkeley National Lab
EC	Environment Canada
ENEA	Italian National Agency of New Technologies, Energy and Environment
FMI	Finnish Meteorological Institute
Harvard	Harvard University
HMS	Hungarian Meteorological Service
IAFMS	Italian Air Force Meteorological Service
JMA	Japan Meteorological Agency
MRI NIES	Meteorological Research Institute, National Institute for Environmental Studies (Japan)
NOAA ESRL	National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Global Monitoring Division (USA)
PNRA	Italian Antarctic Program
SAWS	South African Weather Service
UBA	Umweltsbundesamt Germany
UBAA	Umweltsbundesamt Austria
USDA FS	US Department of Agriculture Forest Service

The data are available from the following sites: NOAA ESRL Observatories: ftp://ftp.cmdl.noaa.gov/ccgg/co2/in-situ/; NOAA ESRL Discrete Surface sites: ftp://ftp.cmdl.noaa.gov/ccgg/co2/flask/event/; NOAA ESRL Aircraft Vertical Profiles from C. Sweeney (NOAA request 3346215); flux tower data at Harvard Forest, Howland Forest and Tapajos from Ameriflux: http://public.ornl.gov/ameriflux/dataproducts.shtml; flux tower data at Northern Old Black Spruce from Harvard University: http://www.seas.harvard.edu/lab/index.html Data Exchange; flux tower data from Southern Great Plains from ARM DAAC: http://www.archive.arm.gov/; Pallas-Sammaltunturni carbon dioxide data from J. Hatakka, personal communication, March 2009; all other observation data from the World Data Center for Greenhouse Gases: http://gaw.kishou.go.jp/wdcgg/wdcgg.html.

A.3. Observation Sites

Observation sites used are presented in **Table A-3** (quasi-continuous), **Table A-4** (discrete surface) and **Table A-5** (discrete aircraft) observations sites. Sites with simulated observations are described in **Table A-6**.

Table **A-3**: Continuous/Quasi-Continuous Observation Sites. Variability indicates the minimum and maximum calculated monthly observation standard deviation in ppm CO_2.

	Location		Variability			
Site	Lat.	Lon.	Min	Max	Agency	Reference
Alert (ALT)	82.45	-62.52	0.50	2.72	EC	Higuchi et al. (2003)
Barrow (BRW)	71.32	-156.60	0.53	3.69	NOAA ESRL	Thoning et al. (2007)
Pallas-Sammaltunturi (PAL)	67.97	24.12	0.74	5.38	FMI	Hatakka et al. (2003), Aalto et al. (2002)
Northern Old Black Spruce (NOBS)	55.88	-98.48	1.21	11.93	Harvard	Dunn et al. (2006)
Fraserdale (FSD)	49.88	-81.57	0.95	5.95	EC	Higuchi et al. (2003)
Schauinsland (SCH)	48.00	8.00	1.24	5.03	UBA	Uhse (2006)
Schneefernerhaus (ZSF)	47.42	10.98	1.06	4.09	UBA	Uhse(2006)
Sonnblick (SNB)	47.05	12.05	1.23	4.06	UBAA	Friedbacher et al. (2007)
Hegyhatsal (HUN)	46.95	16.65	2.35	10.31	HMS	Haszpra et al. (2008)
Plateau Rosa (PRS)	45.93	7.70	0.76	3.44	CESI RICERCA	Apadula et al. (2003)
Park Falls 30m (LEF030)	45.93	-90.27	1.30	8.23	NOAA ESRL	Andrews et al. (2008)
Park Falls 122m (LEF122)	45.93	-90.27	1.28	8.05	NOAA ESRL	Andrews et al. (2008)
Park Falls 396m (LEF396)	45.93	-90.27	1.24	8.11	NOAA ESRL	Andrews et al. (2008)
Howland Forest (HOW)	45.20	-68.74	1.57	8.29	USDA FS	Hollinger et al. (1999)
Monte Cimone (CMN)	44.18	10.70	1.35	3.69	IAFMS	Santaguida (2008)
Harvard Forest (HFM)	42.53	-72.17	2.61	11.16	Harvard	Urbanski et al. (2007)
Ryori (RYO)	39.03	141.82	1.58	8.61	JMA	Esaki (2009)
Southern Great Plains (SGP)	36.80	-97.50	1.66	9.37	DOE LBNL	Fischer (2005), Fischer et al. (2007)
Moody 30m (WKT030)	31.32	-97.33	0.77	6.86	NOAA ESRL	Andrews et al. (2008)
Moody 122m (WKT122)	31.32	-97.33	0.93	6.84	NOAA ESRL	Andrews et al. (2008)
Moody 457m (WKT457)	31.32	-97.33	0.93	6.27	NOAA ESRL	Andrews et al. (2008)
Tenerife (IZO)	28.30	-16.48	0.57	1.78	AEMET	Gomez-Pelaez et al. (2005)
Yonagunijima (YON)	24.47	123.02	0.94	3.78	JMA	Tsutsumi et al. (2006), Esaki (2009)
Minamitorishima (MNM)	24.30	153.98	0.53	2.30	JMA	Esaki (2009)
Mauna Loa (MLO)	19.54	-155.58	0.50	0.99	NOAA ESRL	Thoning et al. (2007)
Tapajos (TPJ)	-2.86	-54.96	1.53	8.28	Harvard	Hutyra et al. (2007)
Tutuila (SMO)	-14.24	-170.57	0.51	1.21	NOAA ESRL	Thoning et al. (2007)
Cape Point (CPT)	-34.35	18.49	0.50	0.99	SAWS	Brunke (2007)
Jubany (JBN)	-62.23	-58.82	0.58	0.98	PNRA, CNR-ICES, DNA	Ciattaglia et al. (1995), Ciattaglia et al. (2008)
South Pole (SPO)	-89.98	-24.80	0.53	0.72	NOAA ESRL	Thoning et al. (2007)

In Table A-4, the discrete surface sites Estevan Point (ESP) and the Pacific Ocean series (POC) are used only in site sensitivity testing. Longitude locations for the Pacific Ocean series are indicative of the model grid sampled and not the observations. See Section 3.n in Chapter 3 for discussion of ESP and Section 4.n in Chapter 4 for the use of the Pacific Ocean series.

Table **A-4**: Discrete Surface Observation Sites. Variability indicates the minimum and maximum calculated monthly observation standard deviation in ppm CO_2.

	Location		Variability			
Site	Lat.	Lon.	Min	Max	Agency	Reference
Alert (ALT)	82.45	-62.52	0.55	3.14	NOAA ESRL	Conway et al. (2008)
Ny-Alesund (ZEP)	80.00	11.88	0.57	3.43	NOAA ESRL	Conway et al. (2008)
Barrow (BRW)	71.32	-156.60	0.50	5.40	NOAA ESRL	Conway et al. (2008)
Station M (STM)	66.00	2.00	0.51	2.98	NOAA ESRL	Conway et al. (2008)
Storhofdi (ICE)	62.00	-20.29	0.53	3.28	NOAA ESRL	Conway et al. (2008)
Baltic Sea (BAL)	55.35	17.22	1.93	9.35	NOAA ESRL	Conway et al. (2008)
Cold Bay (CBA)	55.20	-162.82	0.50	2.96	NOAA ESRL	Conway et al. (2008)
Mace Head (MHD)	53.33	-9.90	0.50	3.54	NOAA ESRL	Conway et al. (2008)
Shemya Island (SHM)	52.72	174.10	0.59	3.79	NOAA ESRL	Conway et al. (2008)
Estevan Point (ESP)	49.38	-126.55	0.64	10.03	EC	Higuchi et al. (2003)
Hegyhatsal (HUN)	46.95	16.65	1.35	10.36	NOAA ESRL	Conway et al. (2008)
Sary Taukum (KZD)	44.45	77.57	0.79	5.27	NOAA ESRL	Conway et al. (2008)
Ulaan Uhl (UUM)	44.45	111.10	0.67	6.41	NOAA ESRL	Conway et al. (2008)
Black Sea (BSC)	44.17	30.00	0.71	12.76	NOAA ESRL	Conway et al. (2008)
Plateau Assy (KZM)	43.25	77.88	0.62	4.51	NOAA ESRL	Conway et al. (2008)
Niwot Ridge (NWR)	40.05	-105.58	0.51	2.95	NOAA ESRL	Conway et al. (2008)
Wendover (UTA)	39.90	-113.72	0.51	3.38	NOAA ESRL	Conway et al. (2008)
Azores	38.77	-27.38	0.50	3.18	NOAA ESRL	Conway et al. (2008)
Tae-Ahn Peninsula (TAP)	36.73	126.13	0.70	7.98	NOAA ESRL	Conway et al. (2008)
Mt. Waliguan (WLG)	36.29	100.90	0.66	3.42	NOAA ESRL	Conway et al. (2008)
Lampedusa (LMP)	35.52	12.62	0.53	4.36	ENEA	Artuso et al. (2009), Apadual et al. (2005), di Sarra et al. (2005)
St. David's Head (BME)	32.37	-64.65	0.51	3.27	NOAA ESRL	Conway et al. (2008)
Tudor Hill (BMW)	32.37	-64.88	0.56	3.46	NOAA ESRL	Conway et al. (2008)
Sede Boker (WIS)	31.13	34.88	0.52	3.71	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN30)	30.00	-126.00	0.52	1.61	NOAA ESRL	Conway et al. (2008)
Tenerife	28.30	-16.48	0.53	1.52	NOAA ESRL	Conway et al. (2008)
Sand Island (MID)	28.21	-177.38	0.51	2.19	NOAA ESRL	Conway et al. (2008)
Key Biscayne (KEY)	25.67	-80.20	0.53	4.48	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN25)	25.00	-134.00	0.63	2.58	NOAA ESRL	Conway et al. (2008)
Assekrem (ASK)	23.18	5.42	0.50	1.20	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN20)	20.00	-141.00	0.62	1.67	NOAA ESRL	Conway et al. (2008)
Mauna Loa (MLO)	19.54	-155.58	0.50	1.26	NOAA ESRL	Conway et al. (2008)
Cape Kumakahi (KUM)	19.52	-154.82	0.56	1.91	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN15)	15.00	-147.00	0.55	1.34	NOAA ESRL	Conway et al. (2008)
Mariana Islands (GMI)	13.43	144.78	0.50	1.38	NOAA ESRL	Conway et al. (2008)
Ragged Point (RPB)	13.17	-59.43	0.52	1.70	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN10)	10.00	-152.00	0.58	1.30	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCN05)	5.00	-158.00	0.52	1.26	NOAA ESRL	Conway et al. (2008)
Christmas Island (CHR)	1.70	-157.17	0.55	0.97	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POC000)	0.00	-163.00	0.50	0.97	NOAA ESRL	Conway et al. (2008)
Mahe Island (SEY)	-4.67	55.17	0.54	1.74	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCS05)	-5.00	-168.00	0.52	0.99	NOAA ESRL	Conway et al. (2008)

Table A-4. (continued)

Site	Location		Variability		Agency	Reference
	Lat.	Lon.	Min.	Max.		
Ascension Island (ASC)	-7.92	-14.42	0.50	1.06	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCS10)	-10.00	-174.00	0.50	1.70	NOAA ESRL	Conway et al. (2008)
Tutuila (SMO)	-14.24	-170.57	0.53	1.49	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCS15)	-15.00	-178.00	0.56	0.95	NOAA ESRL	Conway et al. (2008)
Cape Ferguson (CFA)	-18.00	147.06	0.54	2.71	CSIRO	Francey et al. (2003), Langenfelds et al. (2002)
Pacific Ocean (POCS20)	-20.00	-178.50	0.54	0.88	NOAA ESRL	Conway et al. (2008)
Pacific Ocean (POCS25)	-25.00	174.00	0.51	0.86	NOAA ESRL	Conway et al. (2008)
Cape Grim (CGO)	-42.00	142.50	0.52	0.88	NOAA ESRL	Conway et al. (2008)
Macquarie Island (MQA)	-54.48	159.97	0.51	0.99	CSIRO	Francey et al. (2003), Langenfelds et al. (2002)
Palmer Station (PSA)	-66.00	-64.00	0.51	0.85	NOAA ESRL	Conway et al. (2008)
Mawson (MAA)	-67.62	62.87	0.51	0.84	CSIRO	Francey et al. (2008), Langenfelds et al. (2002)
Syowa Station (SYO)	-70.00	39.58	0.50	0.87	NOAA ESRL	Conway et al. (2008)
Halley Station (HBA)	-75.58	-25.00	0.50	0.81	NOAA ESRL	Conway et al. (2008)
South Pole (SPO)	-89.98	-24.80	0.52	0.77	NOAA ESRL	Conway et al. (2008)

Table **A-5**: Discrete Aircraft Observation Sites. Variability indicates the minimum and maximum calculated monthly observation standard deviation in ppm CO_2.

	Location		Variability			
Site	Lat.	Lon.	Min	Max	Agency	Reference
Briggsdale 3000m (CAR)	40.37	-104.30	0.52	4.49	NOAA ESRL	NOAA, C. Sweeney
Briggsdale 4000m (CAR)	40.37	-104.30	0.53	3.41	NOAA ESRL	NOAA, C. Sweeney
Briggsdale 5000m (CAR)	40.37	-104.30	0.54	3.25	NOAA ESRL	NOAA, C. Sweeney
Briggsdale 6000m (CAR)	40.37	-104.30	0.52	1.82	NOAA ESRL	NOAA, C. Sweeney
Briggsdale 7000m (CAR)	40.37	-104.30	0.54	1.64	NOAA ESRL	NOAA, C. Sweeney
Briggsdale 8000m (CAR)	40.37	-104.30	0.57	2.06	NOAA ESRL	NOAA, C. Sweeney
JAL (WPO)	30.00	146.00	0.57	2.90	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	25.00	146.00	0.56	2.79	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	20.00	146.00	0.57	2.18	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	15.00	146.00	0.54	1.82	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	10.00	146.00	0.50	1.71	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	5.00	146.00	0.50	1.27	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	0.00	146.00	0.54	1.58	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	-5.00	146.00	0.51	1.48	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	-10.00	146.00	0.52	1.05	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	-15.00	146.00	0.54	0.96	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	-20.00	146.00	0.51	1.32	MRI NIES	Matsueda et al. (2008)
JAL (WPO)	-25.00	146.00	0.50	2.04	MRI NIES	Matsueda et al. (2008)

Table **A-6**: Simulated Observation Sites. Type is the observation type: surface (S), elevated (E), 400m tall tower (T). Sampling location is the model sampling location. Variability indicates the minimum and maximum standard deviation of monthly means (ppm CO_2) of model samples for mid-day hours. Reference Site is an existing observation site near the model sampling location.

		Sampling Location		Variability		
Site No.	Type	Lat.	Lon.	Min.	Max.	Reference Site
1	S	54.95	-112.47	1.90	7.57	Lac LaBiche AB (LLB)
2	S	53.87	-104.65	0.94	4.59	Candle Lake SK (CDL)
3	S	45.03	-68.68	0.79	7.32	Argyle ME (AMT)
4	S	44.72	-96.09	2.32	11.48	Rosemount MN
5	S	44.67	-124.07	1.11	4.32	Oregon Transect
6	E	44.50	-123.55	1.04	4.44	Oregon Transect
7	E	44.45	-121.56	1.04	4.42	Oregon Transect
8	E	43.45	-119.72	1.33	4.82	Oregon Transect
9	S	43.93	-60.02	0.60	5.29	Sable Island NS (WSA)
10	T	41.72	-91.35	2.02	11.89	West Branch IA (WBI)
11	E	40.56	-111.64	0.70	2.79	Hidden Peak UT (HDP)
12	E	40.45	-106.73	0.81	2.48	Storm Peak Lab CO (SPL)
13	E	40.05	-105.58	0.81	2.78	Niwot Ridge CO (NWR)
14	T	40.05	-105.01	1.28	4.35	Erie CO (BAO)
15	S	41.48	-96.44	1.99	10.14	Mead NE
16	E	39.91	-105.88	1.07	3.27	Fraser Forest CO (FEF)
17	S	39.32	-86.41	2.98	10.58	Morgan Monroe IN
18	S	39.06	-79.42	3.16	7.72	Canaan Valley WV
19	S	38.74	-92.20	3.08	10.53	Missouri Ozark MO
20	T	38.57	-121.47	1.01	4.70	Bay Area CA (WGC/STR)
21	T	38.53	-78.43	1.90	7.60	Shenandoah VA (SNP)
22	S	36.46	-109.10	0.59	3.17	Roof Butte AZ
23	S	35.93	-84.33	3.06	8.01	Chestnut Ridge TN
24	T	33.40	-81.83	2.34	8.00	South Carolina Tower (SCT)
25	S	49.71	-112.93	1.72	6.06	Lethbridge AB
26	S	49.69	-74.34	2.22	8.12	Chibougamau QC (CHI)
27	T	48.13	-97.98	1.65	10.32	Dahlen ND
28	S	44.32	-79.93	3.45	8.73	Egbert ON (EGB)
29	S	34.25	-89.97	2.37	8.65	Goodwin MS
30	S	31.70	-110.72	0.70	3.18	Audubon Ranch CA
31	S	29.77	-82.20	1.68	5.97	Austin Carey FL
32	S	26.00	-112.50	0.61	1.88	LaPaz, Baja Mexico

A.4. Observation Network Composition

Table **A-7**: Network Reference.

Identifier	Sites	Description
Networks Referenced in Chapter 2		
B	54	Base network consisting of sites from NOAA ESRL, EC, CSIRO.
E	86	Enhanced network: base network with additional sites from WDCGG (World Data Center for Greenhouse Gases).
CE	91	Continental extension network: enhanced network plus 5 continental flux tower sites. This network is used in the Control Inversion.
Networks Referenced in Chapter 3		
CE	91	The 91-site continental extension network from Chapter 2.
N5	92	The CE network with ESP added.
N6	90	The CE network with HFM removed.
N7	90	The CE network with HOW removed.
N8	78	The CE network with no co-located surface sites (sampled in same model grid cell).
Networks Referenced in Chapter 4		
CE	91	The 91-site continental extension network from Chapter 2.
T1	69	The CE network with all continuous continental sites removed.
N1	73	The CE network with all aircraft sites (CAR, WPO) removed.
N2	91	The CE network with POC sites instead of WPO.
N3	86	The CE network with high elevation European sites removed.
N4	88	The CE network with MNM, RYO and YON sites removed.
Networks Referenced in Chapter 5		
CE	91	The 91-site continental extension network from Chapter 2.
V1	79	The CE network with all North American continental continuous sites removed.
V2	86	The V1 network plus the TPJ flux tower, LEF396, WKT457, and the VTT-adjusted observations at HOW, HFM, NOBS and SGP.
V3	87	The V2 network plus FSD.
V4	91	The CE network with surface layer observations at HOW, HFM, NOBS and SGP replaced with VTT-adjusted observations.
F1	115	The CE network plus simulated sites 1-24 from Table A-6 representative of the North American sites active in 2009.
F2	124	The F1 network plus simulated sites 25-32 from Table A-6.

Table **A-8**: Observation Site-Network Cross Reference for Chapter 2. Site Codes and locations of observing sites are in Tables A-3, A-4 and A-5. Networks are identified in Table A-7.

	Networks		
Site Code	B	E	CE
Quasi-Continuous (Table A-3)			
ALT	X	X	X
BRW	X	X	X
PAL		X	X
NOBS			X
FSD	X	X	X
SCH		X	X
ZSF		X	X
SNB		X	X
HUN		X	X
PRS		X	X
LEF030	X	X	X
LEF122	X	X	X
LEF396	X	X	X
HOW			X
CMN		X	X
HFM			X
RYO		X	X
SGP			X
WKT030	X	X	X
WKT122	X	X	X
WKT457	X	X	X
IZO		X	X
YON		X	X
MNM		X	X
MLO	X	X	X
TPJ			X
SMO	X	X	X
CPT		X	X
JBN		X	X
SPO	X	X	X
Discrete Surface (Table A-4)			
All sites except ESP,LMP,POC	X	X	X
LMP		X	X
Discrete Aircraft (Table A-5)			
CAR (all)		X	X
WPO (all)		X	X

Table **A-9**: Observation Site-Network Cross Reference for Chapter 3. Site codes and locations of observing sites are in Tables A-3, A-4 and A-5. Networks are identified in Table A-7.

	Networks				
Site Code	CE	N5	N6	N7	N8
Quasi-Continuous (Table A-3)					
ALT	X	X	X	X	X
BRW	X	X	X	X	X
PAL	X	X	X	X	X
NOBS	X	X	X	X	X
FSD	X	X	X	X	X
SCH	X	X	X	X	X
ZSF	X	X	X	X	X
SNB	X	X	X	X	X
HUN	X	X	X	X	X
PRS	X	X	X	X	X
LEF030	X	X	X	X	
LEF122	X	X	X	X	
LEF396	X	X	X	X	X
HOW	X	X	X		X
CMN	X	X	X	X	X
HFM	X	X		X	X
RYO	X	X	X	X	X
SGP	X	X	X	X	X
WKT030	X	X	X	X	
WKT122	X	X	X	X	
WKT457	X	X	X	X	X
IZO	X	X	X	X	X
YON	X	X	X	X	X
MNM	X	X	X	X	X
MLO	X	X	X	X	X
TPJ	X	X	X	X	X
SMO	X	X	X	X	X
CPT	X	X	X	X	X
JBN	X	X	X	X	X
SPO	X	X	X	X	X
Discrete Surface (Table A-4)					
ALT	X	X	X	X	
ZEP	X	X	X	X	X
BRW	X	X	X	X	
STM	X	X	X	X	X
ICE	X	X	X	X	X
BAL	X	X	X	X	X
CBA	X	X	X	X	X
MHD	X	X	X	X	X
SHM	X	X	X	X	X
ESP		X			
HUN	X	X	X	X	
KZD	X	X	X	X	

Table A-9 (continued)

Site Code	Networks CE	N5	N6	N7	N8
Discrete Surface (continued)					
UUM	X	X	X	X	X
BSC	X	X	X	X	X
KZM	X	X	X	X	X
NWR	X	X	X	X	X
UTA	X	X	X	X	X
AZR	X	X	X	X	X
TAP	X	X	X	X	X
WLG	X	X	X	X	X
LMP	X	X	X	X	X
BME	X	X	X	X	
BMW	X	X	X	X	X
WIS	X	X	X	X	X
IZO	X	X	X	X	
MID	X	X	X	X	X
KEY	X	X	X	X	X
ASK	X	X	X	X	X
MLO	X	X	X	X	
KUM	X	X	X	X	X
GMI	X	X	X	X	X
RPB	X	X	X	X	X
CHR	X	X	X	X	X
SEY	X	X	X	X	X
ASC	X	X	X	X	X
SMO	X	X	X	X	
CFA	X	X	X	X	X
CGO	X	X	X	X	X
MQA	X	X	X	X	X
PSA	X	X	X	X	X
MAA	X	X	X	X	X
SYO	X	X	X	X	X
HBA	X	X	X	X	X
SPO	X	X	X	X	
Discrete Aircraft (Table A-5)					
CAR (all)	X	X	X	X	X
WPO (all)	X	X	X	X	X

Table **A-10**: Observation Site-Network Cross Reference for Chapter 4. Site codes and locations of observing sites are in Tables A-3, A-4, and A-5. Networks are identified in Table A-7.

	Networks					
Site Code	CE	T1	N1	N2	N3	N4
Quasi-Continuous (Table A-3)						
ALT	X	X	X	X	X	X
BRW	X	X	X	X	X	X
PAL	X		X	X	X	X
NOBS	X		X	X	X	X
FSD	X		X	X	X	X
SCH	X		X	X		X
ZSF	X		X	X		X
SNB	X		X	X		X
HUN	X		X	X	X	X
PRS	X		X	X		X
LEF030	X		X	X	X	X
LEF122	X		X	X	X	X
LEF396	X		X	X	X	X
HOW	X		X	X	X	X
CMN	X		X	X		X
HFM	X		X	X	X	X
RYO	X		X	X	X	
SGP	X		X	X	X	X
WKT030	X		X	X	X	X
WKT122	X		X	X	X	X
WKT457	X		X	X	X	X
IZO	X	X	X	X	X	X
YON	X		X	X	X	
MNM	X	X	X	X	X	
MLO	X	X	X	X	X	X
TPJ	X		X	X	X	X
SMO	X	X	X	X	X	X
CPT	X		X	X	X	X
JBN	X	X	X	X	X	X
SPO	X	X	X	X	X	X
Discrete Surface (Table A-4)						
All except ESP, POC	X	X	X	X	X	X
POC (all)				X		
Discrete Aircraft (Table A-5)						
CAR (all)	X	X		X	X	X
WPO (all)	X	X			X	X

Table **A-11**: Observation Site-Network Cross Reference for Chapter 5. Site Codes and locations of observing sites are in Tables A-3, A-4, A-5, and A-6. Networks are identified in Table A-7.

	Networks						
Site Code	CE	V1	V2	V3	V4	F1	F2
Quasi-Continuous (Table A-3)							
ALT	X	X	X	X	X	X	X
BRW	X	X	X	X	X	X	X
PAL	X	X	X	X	X	X	X
NOBS	X					X	X
NOBS VTT			X	X	X		
FSD	X			X	X	X	X
SCH	X	X	X	X	X	X	X
ZSF	X	X	X	X	X	X	X
SNB	X	X	X	X	X	X	X
HUN	X	X	X	X	X	X	X
PRS	X	X	X	X	X	X	X
LEF030	X				X	X	X
LEF122	X				X	X	X
LEF396	X		X	X	X	X	X
HOW	X					X	X
HOW VTT			X	X	X		
CMN	X	X	X	X	X	X	X
HFM	X					X	X
HFM VTT			X	X	X		
RYO	X	X	X	X	X	X	X
SGP	X					X	X
SGP VTT			X	X	X		
WKT030	X				X	X	X
WKT122	X				X	X	X
WKT457	X		X	X	X	X	X
IZO	X	X	X	X	X	X	X
YON	X	X	X	X	X	X	X
MNM	X	X	X	X	X	X	X
MLO	X	X	X	X	X	X	X
TPJ	X		X	X	X	X	X
SMO	X	X	X	X	X	X	X
CPT	X	X	X	X	X	X	X
JBN	X	X	X	X	X	X	X
SPO	X	X	X	X	X	X	X
Discrete Surface (Table A-4)							
All except ESP, POC	X	X	X	X	X	X	X
Discrete Aircraft (Table A-5)							
CAR (all)	X	X	X	X	X	X	X
WPO (all)	X	X	X	X	X	X	X
Simulated Observations (Table A-6)							
Site Numbers 1-24						X	X
Site Numbers 25-32							X

BIBLIOGRAPHY

Aalto, T., Hatakka, J., Paatero, J., Tuovinen, J.-P., Aurela, M., Laurila, T. Holmen, K., Trivett, N. and Viisanen, Y. 2002. Tropospheric carbon dioxide concentrations at a northern boreal site in Finland: Basic variations and source areas. *Tellus* 54B, 110-126.

Andrews, A.E., Koffler, J., Bakwin, P.S., Zhao, C. and Tans, P. 2008. Carbon Dioxide and Carbon Monoxide Dry Air Mole Fractions from the NOAA ESRL Tall Tower Network, 1992-2008. Version 2008-12-11, Path: ftp://ftp.cmdl.noaa.gov/ccg/towers/.

Apadula, F., Gotti, A., Pigini, A., Longhetto, A., Rocchetti, F., Cassardo, C., Ferrarese, S. and Forza, R. 2003. Localization of source and sink regions of carbon dioxide through the method of the synoptic air trajectory statistics. *Atmospheric Environment* 37, 3757-3770, doi:10.1016/S1352-2310(03)00505-3.

Apadula, F., Artuso, F., Chamard, P., De Nile, F., di Sarra, A., Lauria, L., Longhetto, A., Monteleone, F., Piacentino, S., Santaguida, R. and Vannini, C. 2005. The network for background CO_2 measurement in Italy, 12th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques. World Meteorological Organization Global Atmosphere Watch Report no. 161 (WMO TD no. 1275), 173-175.

Artuso, F., Chamard, P., Piacentino, S., Sferlazzo, D., De Silvestri, L., di Sarra, A., Meloni, D. and Monteleone, F. 2009. Influence of transport and trends in atmospheric CO_2 at Lampedusa. *Atmospheric Environment* 43(19), 3044-3051.

Baker, D.F., Law, R.M., Gurney, K.R., Rayner, P., Peylin, P., Denning, A.S., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I.Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Masarie, K., Prather, M., Pak, B., Taguchi, S. and Zhu, Z. 2006. TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988-2003. *Global Biogeochemical Cycles* 20, GB1002, doi:10.1029/2004GB002439.

Baker, I.T., Berry, J.A., Collatz, G.J., Denning, A.S., Hanan, N.P., Philpott, A.W., Prihodko, L., Schaefer, K.M., Stockli, R.S. and Suits, N.S. 2007. Global Net Ecosystem Exchange (NEE) fluxes of CO_2. ORNL Distrib. Active Arch. Cent., Oak Ridge, Tenn., USA. (Available at http://www.daac.ornl.gov)

Baker, I.T., Prihodko, L., Denning, A.S., Goulden, M., Miller, S. and da Rocha, H.R. 2008. Seasonal drought stress in the Amazon: Reconciling models and observations. *Journal of Geophysical Research* 113, G00B01, doi:10.1029/2007JG000644.

Bakwin, P.S., Tans, P.P., Hurst, D.F. and Zhao, C. 1998. Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program. *Tellus* 50B, 401-415.

Bloom, S., da Silva, A. and Dee, D. 2005. Documentation and validation of the Goddard Earth Observing System, Version 4. Technical Report Series on Global Modeling and Data Assimilation, Greenbelt, MD.

Bousquet, P., Peylin, P., Ciais, P., LeQuéré, C., Friedlingstein, P. and Tans, P.P. 2000. Regional changes in carbon dioxide fluxes of land and oceans since 1980 *Science* 290, 1342-1346.

Brenkert, A. L. 1998. Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis. NDP-058A, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., Oak Ridge, Tenn., USA. (Available at http://cdiac.ornl.gov/epubs/ndp/ndp058a/ndp058a.html)

Brunke, E. 2007. Atmospheric CO_2 hourly concentration data, Cape Point Global Atmosphere Watch station. World Data Center for Greenhouse Gases, Japan Meteorol. Agency, Tokyo. (Available at http://gaw.kishou.go.jp/wdcgg/wdcgg.html).

CCSP. 2007. The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research (King, A.W., Dilling, L., Zimmerman. G.P., Fairman, D.M., Houghton, R.A., Marland, G., Rose, A.Z. and Wilbanks, T.J. (eds.)). National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, USA, 242 pp.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., Chevallier, F., De Noblet, N., Friend, A.D., Friedlingstein, P., Grünwald, T., Heinesch, B., Keronen, P., Knohl, A., Krinner, G., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J. M., Papale, P., Pilegaard, K., Rambal, S., Seufert, G., Soussana, J.F., Sanz, M.J., Schulze, E.D., Vesala, T. and Valentini, R. 2005. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* 437, 530-533; doi:10.1038/nature03972.

Ciattaglia, L., Colombo, T. and Masarie, K.A. 1995. Continuous measurements of atmospheric CO_2 at Jubany Station, Antarctica. *Tellus* 51B, 713-721.

Ciattaglia, L., Rafanelli, C., Araujo, J. and Rodriguez, H. 2008. Long-term measurements of the atmospheric carbon dioxide concentration measured at Jubany Station indicate a relationship with "El Niño". *Berichte zur Polar und Meeresforschung*, AWI Bremerhaven, 571-2008, 390-396.

Conway, T.J., Lang, P.M. and Masarie, K.A. 2008. Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2007. Version: 2008-07-24, Path: ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/.

Crisp, D., Atlas, R.M., Breon, F.-M., Brown, L.R., Burrows, J.P., Ciais, P., Connor, B.J., Doney, S.C., Fung, I.Y., Jacob, D.J., Miller, C.E., O'Brien, D., Pawson, S., Randerson, J.T.,

Rayner, P., Salawitch, R.J., Sander, S.P., Sen, B., Stephens, G.L., Tans, P.P., Toon, G.C., Wennberg, P.O., Wofsy, S.C., Yung, Y.L., Kuang, Z., Chaudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D. and Schroll, S. 2004. The Orbiting Carbon Observatory (OCO) mission. *Advances in Space Research* 34(4), 700-709, doi:10.1016/j.asr.2003.08.062.

Davis, K.J., Bakwin, P.S., Yi, C., Berger, B.W., Zhao, C., Teclaw, R.M. and Isebrands, J.G. 2003. The annual cycles of CO_2 and H_2O exchange over a northern mixed forest as observed from a very tall tower. *Global Change Biology* 9, 1278-1293.

Denman, K.L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P.M., Dickinson, R.E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., da Silva Dias, P.L., Wofsy, S.C. and Zhang, X. 2007. Couplings between changes in the climate system and biogeochemistry. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller. pp. 501-588, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Denning, A.S., Fung, I.Y. and Randall, D.A. 1995. Latitudinal gradient of atmospheric CO_2 due to seasonal exchange with land biota. *Nature* 376, 240-243.

di Sarra, A., Chamard, P., Piacentino, S., Monteleone, F., Ciattaglia, L. and Artuso, F. 2005. Influence of the CO_2 latitudinal gradient on the observations at the Mediterranean island of Lampedusa. Seventh International Carbon Dioxide Conference, Extended Abstracts, ISBN 0-9772755-0-7, Published by Committee of Seventh International Carbon Dioxide Conference, National Oceanic and Atmospheric Administration, FF-254 225.

Dunn, A.L., Barford, C.C., Wofsy, S.C., Goulden, M.L. and Daube, B.C. 2006. A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability, and decadal trends. *Global Change Biology* 12, doi:10.1111/j.1365-2486.2006.01221.x.

Engelen, R.J., Denning, A.S., Gurney, K.R., Law, R.M., Rayner, P.J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y-H., Ciais, P., Fan, S., Fung, I.Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B.C., Sarmiento, J., Taguchi, S., Takahashi, T. and Yuen, C.-W. 2002. On error estimation in atmospheric CO_2 inversions. *Journal of Geophysical Research* 107, 4635, doi:10.1029/2002JD002195.

Enting, I.G. 2002. *Inverse problems in atmospheric constituent transport*. 392 pp., Cambridge Univ. Press, Cambridge, UK.

Erickson, D.J., III, Mills, R.T., Gregg, J., Blasing, T.J., Hoffman, F.M., Andres, R.J., Devries, M., Zhu, Z. and Kawa, S.R. 2008. An estimate of monthly global emissions of anthropogenic CO_2: Impact on the seasonal cycle of atmospheric CO_2. *Journal of Geophysical Research* 113, G01023, doi:10.1029/2007JG000435.

Esaki, Y. 2009. Atmospheric CO_2 hourly concentration data, Minamitorishmi, Ryori and Yonagunijima. World Data Center for Greenhouse Gases, Japan Meteorological Agency, Tokyo. (Available at http://gaw.kishou.go.jp/wdcgg/wdcgg.html)

Fischer, M. 2005. Carbon Dioxide Flux Measurement Systems Handbook. Atmospheric Radiation Measurement Tech Report: ARM-TR048.

Fischer, M., Billesbach, D.P., Berry, J.A., Riley, W.J. and Torn, M.S. 2007. Spatiotemporal variations in growing season exchanges of CO_2, H_2O, and sensible heat in agricultural fields of the Southern Great Plains. Lawrence Berkeley National Laboratory Professional Paper: LBNL-63014.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nhanga, J., Prinn, R., Raga, G., Schulz, M. and Van Dorland, R. 2007. Changes in atmospheric constituents and in radiative forcing. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller. pp. 131-234, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Francey, R.J., Steele, L.P., Spencer, D.A., Langenfelds, R.L., Law, R.M., Krummel, P.B., Fraser, P.J., Etheridge, D.M., Derek, N., Coram, S.A., Cooper, L.N., Allison, C.E., Porter, L. and Baly, S. 2003. The CSIRO (Australia) measurement of greenhouse gases in the global atmosphere. In: Report of the 11th WMO/IAEA Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, Tokyo, Japan, September 2001, S. Toru and S. Kazuto (editors), World Meteorological Organization Global Atmosphere Watch, 97-111.

Friedbacher, E., Froehlich, M. and Spangl, W. 2007. Atmospheric CO_2 hourly concentration data, Sonnblick. World Data Center for Greenhouse Gases, Japan Meteorol. Agency, Tokyo. (Available at http://gaw.kishou.go.jp/wdcgg/wdcgg.html).

Gerbig, C., Lin, J.C., Wofsy, S.C., Andrews, A.E., Daube, B.C., Davis, K.J. and Grainger, C.A. 2003. Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *Journal of Geophysical Research* 108, 4757, doi:10.1029/2003JD003770.

Giglio, L., van der Werf, G.R., Randerson, J.T., Collatz, G.J. and Kasibhatla, P.S. 2006. Global estimation of burned area using MODIS active fire observations. *Atmospheric Chemistry and Physics* 6, 957-974.

GLOBALVIEW-CO2. 2007. Cooperative Atmospheric Data Integration Project - Carbon Dioxide. CD-ROM, NOAA ESRL, Boulder, Colorado. [Also available on Internet via anonymous FTP to ftp.cmdl.noaa.gov, Path: /ccg/co2/GLOBALVIEW]

Gloor, M., Fan, S.M., Pacala, S. and Sarmiento, J. 2000. Optimal sampling of the atmosphere for purpose of inverse modeling: A model study. *Global Biogeochemical Cycles* 14, 407-428.

Gomez-Pelaez, A.J., Ramos, R. and Perez-delaPuerta, J. 2006. Methane and carbon dioxide continuous measurements at Izana GAW station (Spain). In: GAW Report (No. 168) of the 13th WMO/IAEA Meeting of Experts on Carbon Dioxide and Related Tracer Measurement Techniques (Boulder, Colorado, USA, 19-22 September 2005), edited by J.B. Miller, World Meteorological Organization (TD No. 1359), 180-184.

Gourdji, S.M., Mueller, K.L., Schaefer, K. and Michalak, A.M. 2008. Global monthly averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data. *Journal of Geophysical Research* 113, D21115, doi:10.1029/2007JD009733.

Gurney, K.R., Law, R., Rayner, P. and Denning, A.S. 2000. TransCom 3 Experimental Protocol. *Paper 707*, Department of Atmospheric Science, Colorado State University, USA. (Available at http://www.purdue.edu/transcom)

Gurney, K.R., Law, R.M., Denning, A.S., Rayner, P.J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y-H., Ciais, P., Fan, S., Fung, I.Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B.C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T. and Yuen, C.-W. 2002. Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models. *Nature* 415, 626-630.

Gurney, K.R., Law, R.M., Denning, A.S., Rayner, P.J., Pak, B.C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y-H., Ciais, P., Fung, I.Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Peylin, P., Prather, M. and Taguchi, S. 2004. Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks *Global Biogeochemical Cycles* 18, GB1010, doi:10.1029/2003GB002111.

Gurney, K.R., Chen, Y-H., Maki, T., Kawa, S.R., Andrews, A. and Zhu, Z. 2005. Sensitivity of atmospheric CO_2 inversions to seasonal and interannual variations in fossil fuel emissions. *Journal of Geophysical Research* 110, D10308, doi:10.1029/2004JD005373.

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

Haszpra, L., Barcza, Z., Hidy, D., Szilágyi, I., Dlugokencky, E. and Tans, P. 2008. Trends and temporal variations of major greenhouse gases at a rural site in Central Europe. *Atmospheric Environment* 42, 8707-8716, doi:10.1016/j.atmosenv.2008.09.012.

Hatakka, J., Aalto, T., Aaltonen, V., Aurela, M., Hakola, H., Komppula, M., Laurila, T., Lihavainen, H., Paatero, J., Salminen, K. and Viisanen, Y. 2003. Overview of the atmospheric research activities and results at Pallas GAW station. *Boreal Environment Research* 8, 365-383.

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

Hollinger, D.Y., Goltz, S.M., Davidson, E.A., Lee, J.T., Tu, K. and Valentine, H.T. 1999. Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest. *Global Change Biology* 5, 891-902.

Hurwitz, M.D., Ricciuto, D.M., Bakwin, P.S., Davis, K.J., Wang, W., Yi, C. and Butler, M.P. 2004. Transport of carbon dioxide in the presence of storm systems over a northern Wisconsin forest. *Journal of the Atmospheric Sciences* 61, 607-618.

Hutyra, L.R., Munger, J.W., Saleska, S.R., Gottlieb, E., Daube, B.C., Dunn, A.L., Amaral, D.F., de Camargo, P.B. and Wofsy, S.C. 2007. Seasonal controls on the exchange of carbon and water in an Amazonian rain forest. *Journal of Geophysical Research* 112, G03008, doi:10.1029/2006JG000365.

Jacobson, A.R., Mikaloff Fletcher, S.E., Gruber, N., Sarmiento, J.L. and Gloor, M. 2007a. A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Global Biogeochemical Cycles* 21, GB1019, doi:10.1029/2005GB002556.

Jacobson, A.R., Mikaloff Fletcher, S.E., Gruber, N., Sarmiento, J.L. and Gloor, M. 2007b. A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 2. Regional results. *Global Biogeochemical Cycles* 21, GB1020, doi:10.1029/2006GB002703.

Kaminski, T., Rayner, P.J., Heimann, M. and Enting, I.G. 2001. On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research* 106(D5), 4703-4715.

Kawa, S.R., Erickson III, D.J., Pawson, S. and Zhu, Z. 2004. Global CO_2 transport simulations using meteorological data from the NASA data assimilation system. *Journal of Geophysical Research* 109, D18312, doi:10.1029/2004JD004554.

Keeling, C.D., Whorf, T.P., Wahlen, M. and van der Pilcht, J. 1995. Internannual extremes in the rate of rise of atmospheric carbon dioxide since 1980. *Nature* 375, 666-670.

Keeling, R.F., Piper, S.C. and Heimann, M. 1996. Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentrations. *Nature* 381, 218-221.

Langenfelds, R.L., Francey, R.J., Pak, B.C., Steele, L.P., Lloyd, J., Trudinger, C.M. and Allison, C.E. 2002. Interannual growth rate variations of atmospheric CO_2 and its $\delta^{13}C$, H_2, CH_4 and CO between 1992 and 1999 linked to biomass burning. *Global Biogeochemical Cycles* 16, 1048, doi:10.1029/2001GB001466.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K.J., Ciais, P., Denning, A.S. and Rayner, P.J. 2008. Mesoscale inversion: First results from the CERES campaign with synthetic data. *Atmospheric Chemistry and Physics* **8**, 3459-3471.

Law, R.M., Rayner, P.J., Steele, L.P. and Enting, I.G. 2003. Data and modelling requirements for CO_2 inversions using high-frequency data. *Tellus* 55B, 512-521.

Law, R.M., Peters, W., Rödenbeck, C., Aulagnier, C., Baker, I., Bergmann, D.J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P.J., Christensen, J.H., Delage, F., Denning,

A.S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S.R., Kleist, J., Krol, M.C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Patra, P.K., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Takigawa, M., Vautard, R., Vermeulen, A.T. and Zhu, Z. 2008. TransCom model simulations of hourly atmospheric CO_2: Experimental overview and diurnal cycle results for 2002. *Global Biogeochemical Cycles* 22, GB3009, doi:10.1029/2007GB003050.

Levinson, D.H. and Waple, A.M. 2004. State of the climate in 2003. *Bulletin of the American Meteorological Society* 85, S1-S72.

Li. Y.-F. 1996. Global population distribution 1990: Terrestrial area and country name information on a one by one degree grid cell basis. *ORNL/CDIAC-96,* DB1016, Carbon Dioxide Inf. Anal. Cent., Oak Ridge, Tenn., USA. (Available at http://cdiac.ornl.gov/ndps/db1016.html)

Marland, G., Boden, T.A. and Andres, R.J. 2007. Global, regional, and national fossil fuel CO_2 emissions. In: *Trends, A Compendium of Data on Global Change*. Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., U.S. Dep. of Energy, Oak Ridge, Tenn., USA.

Masarie, K.A., Langenfelds, R.L., Allison, C.E., Conway, T.J., Dlugokencky, E.J., Francey, R.J., Novelli, P.C., Steele, L.P., Tans, P.P., Vaughn, B. and White, J.W.C. 2001. NOAA/CSIRO Flask Air Intercomparison Experiment: A strategy for directly assessing consistency among atmospheric measurements made by independent laboratories. *Journal of Geophysical Research* 106(D17), 20445-20464.

Matsueda, H., Machida, T., Sawa, Y., Nakagawa, Y., Hirotani, K., Ikeda, H., Kondo, N. and Goto, K. 2008. Evaluation of atmospheric CO_2 measurements from new flask air sampling of JAL airliner observations. *Papers in Meteorology and Geophysics* 59, 1-17, doi:10.2467/mripapers.59.1.

Moeng, C.-H. and Wyngaard, J.C. 1989. Evaluation of turbulent transport and dissipation closures in second-order modeling. *Journal of the Atmospheric Sciences* 46, 2311-2330.

Mueller, K.L., Gourdji, S.M. and Michalak, A.M. 2008. Global monthly averaged CO_2 fluxes using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements. *Journal of Geophysical Research* 113, D21114; doi:10.1029/2007JD009734.

Myneni, R.B., Dong, J., Tucker, C.J., Kaufmann, R.K., Kauppi, P.E., Liski, J., Zhou, L., Alexeyev, V. and Hughes, M.K. 2001. A large carbon sink in the woody biomass of Northern forests *Proceedings of the National Academy of Sciences USA* 98, 14784-14789, doi:10.1073/pnas.261555198.

Nemani, R.R., Keeling, C.D., Hashimoto, H., Jolly, W.M., Piper, S.C., Tucker, C.J., Myneni, R.B. and Running, S.W. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* 300, 1560-1563, doi:10.1126/science.1082750.

Parazoo, N.C., Denning, A.S., Kawa, S.R., Corbin, K.D., Lokupitiya, R.S. and Baker, I.T. 2008. Mechanisms for synoptic variations of atmospheric CO_2 in North America, South America and Europe. *Atmospheric Chemistry and Physics* 8, 7239-7254.

Patra, P.K. and Maksyutov, S. 2003. Sensitivity of optimal extension of CO2 observation networks to model transport. *Tellus* 55B, 498-511.

Patra, P.K., Gurney, K.R., Denning, A.S., Maksyutov, S., Nakazawa, T., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.H., Ciais, P., Fan, S., Fung, I.Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Law, R.M., Maki, T., Pak, B.C., Peylin, P., Prather, M., Rayner, P.J., Sarmiento, J., Taguchi, S., Takahashi, T. and Yuen, C.W. 2006. Sensitivity of inverse estimation of annual mean CO_2 sources and sinks to ocean-only sites versus all-sites observational networks. *Geophysical Research Letters*, **33,** L05814, doi:10.1029/2005GL025403.

Patra, P.K., Law, R.M., Peters, W., Rödenbeck, C., Takigawa, M., Aulagnier, C., Baker, I., Bergmann, D.J., Bousquet, P., Brandt, J., Bruhwiler, L., Cameron-Smith, P.J., Christensen, J.H., Delage, F., Denning, A.S., Fan, S., Geels, C., Houweling, S., Imasu, R., Karstens, U., Kawa, S.R., Kleist, J., Krol, M.C., Lin, S.-J., Lokupitiya, R., Maki, T., Maksyutov, S., Niwa, Y., Onishi, R., Parazoo, N., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R., Vermeulen, A. and Zhu, Z. 2008. TransCom model simulations of hourly atmospheric CO_2: Analysis of synoptic-scale variations for the period 2002-2003. *Global Biogeochemical Cycles* 22, GB4013, doi:10.1029/2007GB003081.

Peters, W., Jacobson, A.R., Sweeney, C., Andrews, A.E., Conway, T.J., Masarie, K., Miller, J.B., Bruhwiler, L.M.P., Pétron, G., Hirsch, A.I., Worthy, D.E.J., van der Werf, G.R., Randerson, J.T., Wennberg, P.O., Krol, M.C. and Tans, P.P. 2007. An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences USA* 104, 18925-18930, doi:10.1073/pnas.0708986104.

Peylin, P., Baker, D., Sarmiento, J., Ciais, P. and Bousquet, P. 2002. Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data. *Journal of Geophysical Research* 107, D19, 4385, doi:10.1029/2001JD000857.

Peylin, P., Rayner, P.J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., Adolphsen, A., Apadula, F., Graul, R., Haszpra, L., Ramonet, M., Santaguida, R. and Schmidt, M. 2005. Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology. *Atmospheric Chemistry and Physics* 5, 3173-3186.

Potter, C., Klooster, S., Steinbach, M., Tan, P., Kumar, V., Shekhar, S., Nemani, R. and Myneni, R. 2003. Global teleconnections of climate to terrestrial carbon flux. *Journal of Geophysical Research* 108, 4556; doi:10.1029/2002JD002979.

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 Biogeochemical Cycles* 11, 535-560.

Rayner, P.J. 2004. Optimizing CO_2 observing networks in the presence of model error: results from TransCom3. *Atmospheric Chemistry and Physics* 4, 413-421.

Rayner, P.J., Enting, I.G., Francey, R.J. and Langenfelds, R. 1999. Reconstructing the recent carbon cycle from atmospheric CO_2, $\delta^{13}C$ and O_2/N_2 observations. *Tellus* 51B, 213-232.

Rödenbeck, C., Houweling, S., Gloor. M. and Heimann, M. 2003a. CO_2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics* 3, 1919-1964.

Rödenbeck, C., Houweling, S., Gloor, M. and Heimann, M. 2003b. Time-dependent atmospheric CO_2 inversions based on interannually varying tracer transport. *Tellus* 55B, 488-497.

Santaguida, R. 2008. Atmospheric CO_2 hourly concentration data, Monte Cimone. World Data Center for Greenhouse Gases, Japan Meteorol. Agency, Tokyo. (Available at http://gaw.kishou.go.jp/wdcgg/wdcgg.html).

Schuh, A.E., Denning, A.S., Uliasz, M. and Corbin, K.D. 2009. Seeing the forest for the trees: Recovering large-scale carbon flux biases in the midst of small-scale variability. *Journal of Geophysical Research* 114, G03007, doi:10.1029/2008JG000842.

Shumway, R.H. and Stoffer, D.S. 2006. *Time Series Analysis and Its Applications With R Examples*. 2nd ed., 575 pp., Springer, New York, NY, USA.

Stephens, B.B., Gurney, K.G., Tans, P.P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R.L., Steele, L.P., Francey, R.J. and Denning, A.S. 2007. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2. *Science* 316, doi:10.1126/science.1137004.

Suntharalingam, P., Randerson, J.T., Krakauer, N., Logan, J.A. and Jacob, D. J. 2005. Influence of reduced carbon emissions and oxidation on the distribution of atmospheric CO_2: Implications for inverse analyses. *Global Biogeochemical Cycles* 19, GB4003, doi:10.1029/2005GB002466.

Takahashi, T., Sutherland, S.C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., Bates, N., Wanninkhof, R., Feely, R.A., Sabine, C.L., 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 Research II* 49, 1601-1622.

Takahashi, T., Sutherland, S.C., Wanninkhof, R., Sweeney, C., Feely, R.A., Chipman, D.W., Hales, B., Friederich, G.E., Chavez, F.P., Sabine, C.L., Watson, A.J., Bakker, D.C.E., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Kortzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T.S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C.S., Delille, B., Bates, N.R. and de Baar, H.J.W. 2009. Climatological mean and decadal change in surface ocean pCO_2, and

net sea-air CO_2 flux over the global oceans. *Deep-Sea Research II* 56, 554-577, doi:10.1016/j.dsr2.2008.12.009.

Tans, P.P., Fung, I.Y. and Takahashi, T. 1990. Observational constraints on the global atmospheric CO_2 budget. *Science* 247, 1431-1438.

Tans, P., Zhao, C. and Masarie, K. 2003. Maintenance and propagation of the WMO Mole Fraction Scale for carbon dioxide in air. In: Report of the Eleventh WMO Meeting of Experts on Carbon Dioxide Concentration and Related Tracer Measurement Techniques, Tokyo, Japan, 25-28 September 2001, edited by S. Toru, World Meteorological Organization, Geneva.

Tarantola, A. 2005. *Inverse problem theory and methods for model parameter estimation*. 342 pp., Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Taylor, K.E. 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research* 106, D7, doi:10.1029/2000JD900719.

Thoning, K.W., Kitzis, D.R. and Crotwell, A. 2007. Atmospheric Carbon Dioxide Dry Air Mole Fractions from quasi-continuous measurements at Barrow, Alaska; Mauna Loa, Hawaii; American Samoa; and South Pole, 1973-2006. Version: 2007-10-01, Path: ftp://ftp.cmdl.noaa.gov/ccg/co2/in-situ/.

Tsutsumi, Y., Mori, K., Ikegami, M., Tashiro, T. and Tsuboi, K. 2006. Long-term trends of greenhouse gases in regional and background events observed during 1998-2004 at Yonagunijima located to the east of the Asia continent. *Atmospheric Environment* 40, 5868-5879, doi:10.1016/j.atmosenv.2006.04.036.

Uhse, K. 2006. Atmospheric CO_2 hourly concentration data, Schauinsland, Zugspitze/Schneefernerhaus. World Data Center for Greenhouse Gases, Japan Meteorol. Agency, Tokyo. (Available at http://gaw.kishou.go.jp/wdcgg/wdcgg.html).

Urbanski, S., Barford, C., Wofsy, S., Kucharik, C., Pyle, E., Budney, J., McKain, K., Fitzjarrald, D., Czikowsky, M. and Munger, J.W. 2007. Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest. *Journal of Geophysical Research* 112, G02020, doi:10.1029/2006JG000293.

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

Wang, J.-W., Denning, A.S., Lu, L., Baker, I.T., Corbin, K.D. and Davis, K.J. 2007. Observations and simulations of synoptic, regional, and local variations in atmospheric CO_2. *Journal of Geophysical Research* 112, D04108, doi:10.1029/2006JD007410.

Wang, W., Davis, K.J., Yi, C., Patton, E.G., Butler, M.P., Ricciuto, D.M. and Bakwin, P.S. 2007. A note on the top-down and bottom-up gradient functions over a forested site. *Boundary-Layer Meteorology* 124, 305-314, doi:10.1007/s10546-007-9162-0.

Waple, A.M., Lawrimore, J.H., Alexander, L., Horton, B., Ambenje, P., Bell, G.D., Halpert, M.S., Bulygina, O.N., Korshunova, N.N., Razuvaev, V.N., Camargo, S.J., Lyon, B., Seth, A., Christy, J.R., Douglas, A.V., Gleason, K.L., Menne, M.J., Hardy, D., McBride, C., Phillips, D., Pielke, R.S., Castro, C., Robinson, D.A., Salinger, A.J., Schnell, R.C. and Watkins, A. 2003. State of the climate in 2002. *Bulletin of the American Meteorological Society* 84, S1-S68.

Washenfelder, R.A., Toon, G.C., Blavier, J.-F., Yang, Z., Allen, N.T., Wennberg, P.O., Vay, S.A., Matross, D.M. and Daube, B.C. 2006. Carbon dioxide column abundances at the Wisconsin Tall Tower site. *Journal of Geophysical Research* 111**,** D22305, doi:10.1029/2006JD007154.

Wyngaard, J.C. and Brost, R.A. 1984. Top-down and bottom-up diffusion of a scalar in the convective boundary layer. *Journal of the Atmospheric Sciences* 41, 102-112.

Yang, Z., Washenfelder, R.A., Keppel-Aleks, G., Krakauer, N.Y., Randerson, J.T., Tans, P.P., Sweeney, C. and Wennberg, P.O. 2007. New constraints on Northern Hemisphere growing season net flux. *Geophysical Research Letters*34**,** L12807, doi:10.1029/2007GL029742.

Yi, C., Davis, K.J., Berger, B.W. and Bakwin, P.S. 2001. Long-term observations of the dynamics of the continental boundary layer. *Journal of the Atmospheric Sciences* 58, 1288-1299.

Yakota, T., Oguma, H., Morino, I., Higurashi, A., Aoki, T. and Inoue, G. 2004. Test measurements by a BBM of the nadir-looking SWIR FTS aboard GOSAT to monitor CO2 column density from space. *Proceedings of SPIE* 5652, 182, doi:10.1117/12.578497.

Zhao, C.L., Tans, P.P. and Thoning, K.W. 1997. A high precision manometric system for absolute calibrations of CO_2 in dry air. *Journal of Geophysical Research* 102, 5885-5894.

Zhao, C. and Tans, P.P. 2006. Estimating uncertainty of the WMO Mole Fraction Scale for carbon dioxide in air. *Journal of Geophysical Research* 111, D08S09, doi:10.1029/2005JD006003.

VITA

Martha Peirce Butler

Born in Arlington, Massachusetts, and growing up in Lexington, Massachusetts and York Harbor, Maine, Martha Butler has maintained a strong connection to New England. She is a Phi Beta Kappa graduate of Tufts University (1970) with a major in Chemistry and Mathematics. After a two-year stint Out West during which she managed the Chemistry stockroom at Northern Arizona University, she returned to New England and joined a computer programming class at a major Boston insurance company in 1972. After 5 years of providing programming and systems analysis support for the group insurance sector, she returned to Maine in 1977 to continue her information technology career at a prominent mail order retailer in Freeport, Maine. Twenty-two years later, she left the now much-larger company to pursue a life-long dream to become a scientist. That her focus of study would be as broad-reaching as the global carbon cycle is also consistent with her belief that everything is connected.