Data Assimilation for Atmospheric CO_2: Towards Improved Estimates of CO_2 Concentrations and Fluxes

by

Abhishek Chatterjee

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Environmental Engineering)
in the University of Michigan
2012

Doctoral Committee:

Associate Professor Anna M. Michalak, Chair
Professor Peter Adriaens
Richard J. Engelen, European Centre for Medium Range Weather Forecasts (ECMWF)
Assistant Professor Derek J. Posselt
Professor Richard B. Rood

ACKNOWLEDGEMENTS

*The variations in data assimilation branch out so quickly that it is impossible to cover them all. On more than one occasion, I have succumbed to this tree of variations and consequently, ended up in a muddled and chaotic state of mind that would simply refuse to clear. Such moments are best summarized by Mikhail Tal's statement -"Oh, what a task so harsh/ To drag a hippo from a marsh!"**

I am indebted to my advisor Anna Michalak for repeatedly dragging my mind out of the marsh, and making the science clearer and simpler. It would have been hopeless, quite frankly impossible, to put together any semblance of this work without Anna's guidance and insights into the science questions. Over the last six years, I have received an abundance of ideas, mathematical knowledge and insightful criticisms from her. Anna's chief forte (among her many other strengths!) is her perseverance and passion in doing research. Ultimately, this has always been, and shall be a huge source of inspiration to me. I am thankful to her for introducing me to her science community, and also in helping me get started during the early stages of my graduate career, when I had a rather deluded definition of academic research. Truly it has been a privilege to have studied under her, and I could not have asked for a better advisor.

I thank Richard Rood, Peter Adriaens, Derek Posselt and Richard Engelen, for agreeing to serve on my committee and providing me with valuable insights throughout the course of this work. They helped to keep my PhD on track, and have always been available to discuss a variety of science and research questions. I consider myself fortunate to have had such a distinguished and supportive committee.

I thank Jeffrey Anderson for sharing his expertise on ensemble filters and more importantly, teaching me how to appreciate data assimilation. It is largely due to my interactions

[*] *Excerpt from the 'The Life and Games of Mikhail Tal' but originally based on a couplet from Chukovsky.*

with him over the last couple of years, that I started viewing data assimilation more as an art, and less as a cyclopean mathematical black box. It has been this change in perspective that made this work all the more enjoyable.

I am grateful to David Baker for engaging me in several fruitful discussions on 4D-VAR and even sharing his study notes from his graduate student days! David's impromptu lectures on understanding 4D-VAR from a control theory standpoint (during a drive from Boulder to Fort Collins) and the operational implementation of 4D-VAR for the CO_2 flux estimation problem (while walking in circles around Moscone South), are among the best discourses I have had on variational systems. I have to admit though that these conversations also made me realize (much to my chagrin!) the futility of my attempts in trying to setup a 4D-VAR system, capable of ingesting real satellite data.

During the course of this work, I have had the opportunity to interact with several scientists and data assimilation experts. I thank Andy Jacobson, Peter Rayner, Richard Engelen and Doug Nychka for hosting me and facilitating collaborations with several researchers at their institutions. I owe a debt of gratitude to Carla Cardinali, Frederic Chevallier, Ron Errico, Mike Fisher, Peter Houtekamer, Andrew Lorenc, Wouter Peters, Florence Rabier, Chris Snyder, Olivier Talagrand, Peter Jan van Leeuwen and Jeffrey Whitaker for demystifying the different varieties and flavors of data assimilation. I have had the good fortune to participate in two excellent summer schools: (a) 2009 JCSDA Summer Colloquium on Data Assimilation, Stevenson, USA (http://www.jcsda.noaa.gov/meetings_2009SummerColloq.php), and (b) 2012 International Summer School on Advanced Data Assimilation for Geosciences, Les Houches, France (http://houches2012.gforge.inria.fr/). I thank the organizers of this school in providing me the opportunity to learn from different experts in the field as well as forge new collaborations with my peers.

The simulations and results presented in this dissertation would have been impossible without the help of several data providers and modelers within the CO_2 community. I would like to thank the following scientists for sharing their data that have been an integral part of this dissertation: Arlyn Andrews, Thomas Blumenstock , Chris O'Dell, Ken Davis, Nicholas Deutscher, Danilo Dragoni, Marc Fischer, Mathias Goeckede, David Griffith, Frank Hase, Ralph Keeling, Esko Kyrö, Bev Law, Natasha Miles, Isamu Morino, Bill Munger, Justus Notholt, Matt

Parker, Scott Richardson, Vanessa Sherlock, Britt Stephens, Kimberly Strong, Ralf Sussmann, Colm Sweeney, Thorsten Warneke, Paul Wennberg, Steven Wofsy and Douglas Worthy. The contributions of the ACOS/OCO-2 team in generating the GOSAT-ACOS data and the CalTech team in maintaining the TCCON Data Archive are gratefully acknowledged.

I am grateful to all members of the Michalak Research Group (PUORG), especially Sharon Gourdji and Kim Mueller, for providing important inputs on the science aspects of this work. Vineet Yadav along with Yuping Liu (NASA-GSFC) played a key role in setting up the PCTM transport model, without which assimilation of the GOSAT data would not have been possible.

A special mention must be made of the funding agencies/programs that supported the various components of this work – the NASA Earth and Space Science Fellowship Program, the Rackham Graduate School, the Carnegie Institution for Science, and the National Center for Atmospheric Research.

Finally, portions of this dissertation are adapted from the following manuscripts in press, or in preparation:

- Chatterjee, A., A. M. Michalak, J. L. Anderson, K. L. Mueller, V. Yadav (in press), Towards reliable ensemble Kalman filter estimates of CO_2 fluxes, *Journal of Geophysical Research – Atmospheres*, doi:10.1029/2012JD018176
- Chatterjee, A, R. J. Engelen, A. M. Michalak, S. R. Kawa, C. Sweeney (in prep.), Background error covariance for atmospheric CO_2 data assimilation, in preparation for submission to *Journal of Geophysical Research-Atmospheres*
- Chatterjee, A. and A. M. Michalak (in prep.), Inter-comparison of ensemble and variational data assimilation in the context of a CO_2 source-sink estimation problem, in preparation for submission to *Geoscientific Model Development*

The co-authors of these manuscripts are gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

Data Assimilation for Atmospheric CO_2: Towards Improved Estimates of CO_2 Concentrations and Fluxes

by

Abhishek Chatterjee

Chair: Anna M. Michalak

The lack of a process-level understanding of the carbon cycle is a major contributor to our uncertainty in understanding future changes in the carbon cycle and its interplay with the climate system. Recent initiatives to reduce this uncertainty, including increases in data density and the estimation of emissions and uptake (a.k.a. fluxes) at fine spatiotemporal scales, presents computational challenges that call for numerically-efficient schemes. Often based on data assimilation (DA) approaches, these schemes are common within the numerical weather prediction community.

The goal of this research is to identify fundamental gaps in our knowledge regarding the precision and accuracy of DA for CO_2 applications, and develop suitable methods to fill these gaps. First, a new tool for characterizing background error statistics based on predictions from carbon flux and atmospheric transport models is shown to yield improved estimates of CO_2 concentration fields within an operational DA system at the European Centre for Medium-Range

Weather Forecasts (ECMWF). Second, the impact of numerical approximations within existing DA approaches is explored using a simplified flux estimation problem. It is found that a complex interplay between the underlying numerical approximations and the observational characteristics regulates the performance of the DA methods. Third, a novel and versatile DA method called the geostatistical ensemble square root filter (GEnSRF) is developed to leverage the information content of atmospheric CO_2 observations. The ability of GEnSRF to match the performance of a more traditional inverse modeling approach is confirmed using a series of synthetic data experiments over North America. Fourth, GEnSRF is used to assimilate high-density satellite observations from the recently launched GOSAT satellite, and deliver global data-driven estimates of fine-scale CO_2 fluxes. Diagnostics tools are used to evaluate the benefit of satellite observations in constraining global surface fluxes, relative to a traditional surface monitoring network. Overall, this research has developed, applied, and evaluated a novel set of tools with unique capabilities that increase the credibility of DA methods for atmospheric CO_2 applications. Such advancements are necessary if we are to accurately understand the critical controls over the atmospheric CO_2 growth, and improve our understanding of carbon-climate feedbacks.

Chapter 1

Introduction

Data assimilation is "*an analysis technique in which the observed information is accumulated into the background state by taking advantage of consistency constraints with laws of time evolution and physical properties*" (*Bouttier and Courtier* [1999]). Although this definition is merely a decade old, its roots can be traced back to the late eighteenth century when both Gauss and Legendre are credited with simultaneously discovering the core principles behind data assimilation (DA). While Gauss and Legendre's efforts were aimed at calculating the orbits of heavenly bodies, the mathematical basis of their work paved the way for the least squares method, which forms the most basic concept in DA. Through the last two centuries, Gauss and Legendre's ideas and concepts have been advanced and refined, but the fundamental tenet of minimization of the squared departure between an estimate, and observations and/or background information have remained unchanged. Today DA is more generally cast in a probabilistic framework that formalizes the conjunction of the two states of information (e.g. *Tarantola* [2005]), and the least squares method is derived as a special case of this much more general probabilistic framework.

The popularity and recognition of DA has primarily stemmed from its application to weather forecasting. Post-WWII, a rich scientific period ensued, during which advances in atmospheric general circulation models, weather satellites, and computational infrastructure, culminated in the development of the four dimensional view of DA (or, 4D-VAR – *Sasaki*

[1970a, b, c]). Sasaki's trilogy of papers in the *Monthly Weather Review* and his expansive approach was instrumental in demonstrating the value of DA to a broader scientific audience. Based on Sasaki's and subsequently, the work of *Lorenc* [1986] and *Le Dimet and Talagrand* [1986], the main advantages of DA can be summarized as: 1) the inherent ability to bring disparate sources of information (i.e., models, various sources of data) to achieve the best analysis, with the analysis being 'better' than the individual pieces of information alone, 2) the ability to predict future analysis states, and 3) the ability to ambitiously match available computational resources, even with the inclusion of more physical processes in the model or more complex data streams. It is no surprise then, that over the last several decades, DA tools and application have become ubiquitous in almost all Earth Sciences disciplines (e.g. atmospheric chemistry and air quality, hydrology, oceanography etc.).

More recently, within the last seven years, data assimilation has found increasing usage within the carbon science community (*Rayner* [2010]). Its application within carbon cycle research, however, has been in two separate contexts.

In an *assimilation* context, DA has been applied to generate consistent 4-dimensional fields of atmospheric CO_2 concentrations (e.g. *Engelen et al*. [2009]; *Liu et al*. [2012]) or to estimate parameters of biogeochemical models (e.g. *Rayner et al*. [2005]). In this sense, the aim of the carbon DA system is to integrate together atmospheric, terrestrial and oceanic data together, along with underlying dynamical constraints into a common analysis framework. Applications that make use of the assimilation framework are built on the premise of carbon cycle model development (e.g. *Chatterjee et al*. [in prep.]; *Koffi et al*. [2012]; *Ziehn et al*. [2011b]) and/or the predictive properties of the carbon system (e.g. *Rayner et al*. [2011]; *Ziehn et al*. [2011a]).

Within carbon science, however, DA has gained more popularity in an *inversion* context for inference of CO_2 sources and sinks using atmospheric CO_2 measurements (e.g. *Chevallier et al*. [2005b]; *Rödenbeck et al*. [2005]; *Peters et al*. [2005]; *Baker et al*. [2006a]; *Zupanski et al*. [2007a]; *Lokupitiya et al*. [2008]; *Feng et al*. [2009]; *Gerber and Joos* [2010]; *Miyazaki et al*. [2011]; *Kang et al*. [2011]; *Chatterjee et al*. [in press]). This application is based on an inverse modeling paradigm, in which the basic premise is that given a set of atmospheric CO_2 observations, and using a model of atmospheric transport, it is possible to infer information on the distribution of CO_2 fluxes at the surface of the Earth (e.g. *Enting* [2002]). Historically this problem has been solved in a "batch" mode, in which the inversion is performed in one step by solving a single system of linear equations relating the CO_2 fluxes and the CO_2 measurements (e.g. *Ciais et al*. [2010b]). The batch solution however requires an atmospheric transport model to be run either once per estimated flux region/period combination, or once per observation. This becomes prohibitively expensive as the number of CO_2 measurements and/or the number of fluxes to solve for increase. DA methods are able to minimize the number of atmospheric transport model runs by virtue of their underlying numerical approximations and sidestep the computational challenges. Thus, application to the CO_2 flux estimation problem simply intends to take advantage of the computational efficiency of a DA framework.

In reality, within the inversion context the word 'assimilation' is loosely used since there is no explicit dynamical model against which the observations are assimilated. Or in purely DA terms, there is no forecast step to propagate observation information in time and space. The forecast step is a key component in DA, in which the dynamical model is used to provide a first guess of the state vector before observations are assimilated to update the state vector. By propagating the state vector between different assimilation time steps, the dynamical model

directly contributes to the growth of the eigenvalue spectrum of the state covariance matrix in certain preferred directions and decay in others (*Bengtsson et al*. [2003]; *Furrer and Bengtsson* [2007]). For the CO_2 flux estimation problem, however, there is no applicable dynamical model (e.g. *Peters et al*. [2005]; *Miyazaki et al*. [2011]; *Chatterjee et al*. [in press]) that can evolve the fluxes forward in time. As a consequence, changes to the background error structure are only possible via the information provided by the observations. Such a formulation not only loses out on the predictive capabilities associated with a DA framework, but likely results in the DA approaches performing sub-optimally, if there is a lack of measurements to constrain the problem.

Irrespective of these drawbacks, it is well-accepted at this point that as carbon science becomes increasingly data rich, DA provides the optimal and most computationally efficient framework to process the high-density data for estimating CO_2 sources and sinks. Extracting the maximum information content from the data to obtain realistic flux estimates and uncertainties will allow us to address a variety of science and policy questions that could not be answered before (e.g. *Baker et al*. [2006a]). For example, accurately resolving the surface fluxes at fine spatial and temporal scales will help in quantifying the relative importance of the driving processes at these levels. Again as noted by *Dilling et al*. [2003] working out the precise CO_2 fluxes to the level of individual countries may pave the way for improved carbon management policies and for the verification of international emission treaties. DA is envisaged to play a major role in carbon cycle science over the next decade and has been identified as an important modeling tool in several strategy reports (e.g. European Commission Report, ed. *Schulze et al*. [2009]; U.S. Climate Change Science Program [2009]; U.S. Carbon Cycle Science Plan, ed. *Michalak et al*. [2011]).

Almost all of these strategy reports have also challenged the CO_2-DA community to: a) make existing DA systems more reliable and accurate in their analysis estimates (for e.g., determining the magnitude and distribution of CO_2 sources and sinks), and b) extend existing DA systems to incorporate models of major carbon cycle components and pursue multi-species carbon data assimilation (for e.g., transition from an *inversion* to an *assimilation* framework). Tackling the second issue is technically more difficult and has little scientific value until and unless the first issue is completely addressed. But this also raises a different science question, that *in spite of* CO_2*-DA being in vogue for nearly a decade, why are the analyses (e.g. flux) estimates from DA not deemed reliable or accurate*? While there is no single answer to this question, several reasons can be identified as to why DA has failed to provide robust analyses for carbon science applications.

First and foremost, most of the DA tools that have been applied for CO_2 applications have been borrowed from the numerical weather prediction (NWP) community. NWP-DA is designed to work at short time scales due to the chaotic nature of the underlying process. Conversely CO_2 is a long-lived gas that gets well-mixed in the atmosphere at the scale of months (*Bruhwiler et al*. [2005]). Different operational time-scales means that DA tools borrowed from NWP need to be modified to suit the features of the carbon cycle. Additionally, as stated earlier, specifically for CO_2-DA inversions there is no explicit dynamical model that evolves the estimated fluxes forward in time. The lack of a dynamical model represents a valuable loss of information to the DA system. For the CO_2 inverse problem, the absence of such a model coupled with the highly under-determined nature of the estimation problem (i.e., sparse observation datasets) likely hampers the performance of DA approaches, especially when the analysis is conducted at fine spatial and temporal scales.

Secondly, DA aims to optimally combine the information from the observations as well as some background based on their respective uncertainty estimates. Due to the under-constrained and ill-conditioned nature of the CO_2 flux estimation problem, the background plays a significant role in determining the final analysis. Ill-conditioned in this respect means that several different solutions (and not one unique solution) exist that are equally consistent with the available measurement data (e.g. *Tarantola* [2005]). In an ill-conditioned problem, mass balance constraints imply that estimates recovered in regions under-constrained by the measurements are likely to be unrealistic and compensating for limited information in better-constrained regions. Therefore, some sort of regularization technique is needed to stabilize the solution to these inverse problems that are then provided by the background information. Currently, all existing carbon data assimilation systems aim to provide this background information by specifying an initial estimate of CO_2 fluxes, which are obtained either from biospheric models (for e.g., Carnegie-Ames-Stanford Approach (CASA) model, *Randerson et al*. [1997]) and/or inventories (e.g. *Brenkert* [1998]; *Olivier and Berdowski* [2001]; *Takahashi et al*. [2002]). These estimates are often termed as "bottom-up", as these are scaled up from a process-based understanding of CO_2 surface exchange at plot-level or laboratory scales. Bottom-up estimates depend heavily on simplifying assumptions and have been shown to differ considerably in their estimates of fluxes over large regions (e.g. *Huntzinger et al*. [2012]). Given that *a posteriori* flux estimates represent a compromise between the information contained in the measurements versus the prior fluxes, the posterior flux estimates in under-constrained regions can strongly reflect the characteristics of the specific bottom-up estimate used in the inversion (e.g. *Peters et al*. [2007; 2010]). As a consequence, the final flux estimate can no longer be considered as an independent validation tool for process-based understanding of the carbon cycle.

Thirdly, almost all of the DA methods have some form of underlying numerical approximations that make them computationally tractable. The analysis is based on making several compromises between the computational cost, statistical optimality and physical realism of the assimilation problem. The ill-conditioned nature of the CO_2 flux estimation problem, however, may make the errors due to the numerical approximation grow. None of the CO_2-DA studies to date have attempted to test the impact of these numerical approximations. It is still not clear whether these numerically approximate methods can serve as a suitable long-term replacement of the batch technique in producing robust estimates of CO_2 sources and sinks.

The above-described shortcomings have directly contributed to the uncertainty surrounding the applicability and accuracy of DA approaches for atmospheric CO_2 applications. These shortcomings can be restated in the form of three specific research questions that need to be addressed to increase the credibility of the CO_2-DA systems:

(Q1) What are the modifications necessary to adapt existing DA methods for carbon science?

(Q2) What are the errors incurred due to the numerical approximations in the DA methods, and how do these impact the analysis estimates, and the associated uncertainties?

(Q3) Is there a way of providing the background information in a data assimilation system that would make the posterior analysis less reliant on a single bottom-up estimate?

The core of this dissertation revolves around finding the answers to these questions. As illustrated in Figure 1.1, each of the four objectives investigates one (or more) of these questions in detail. These specific objectives and their relevance to the questions outlined in (Q1), (Q2) and (Q3) are discussed next.

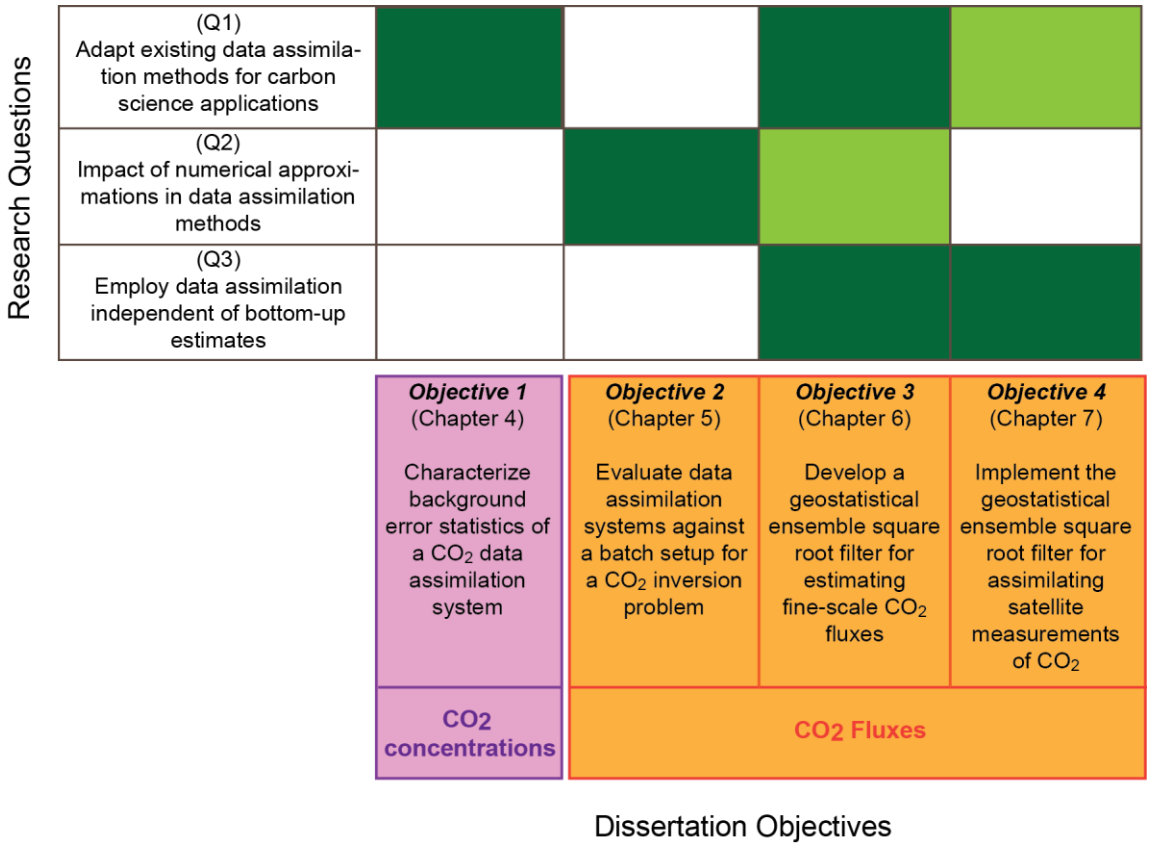


Figure 1.1 *– A schematic showing the correspondence between the three research questions outlined in (Q1) to (Q3) and the four dissertation objectives. Dark green indicates that an objective primarily answers the specific research question. Light green indicates that the analysis conducted as part of the objective helps in answering a secondary research question. In addition, going from Objective 2 to Objective 4, the complexity of the inverse problem that will be solved with DA will increase.*

Objective 1: Characterize the background error statistics for atmospheric CO_2 data assimilation

The first objective (Chapter 4; *Chatterjee et al*. [in prep.]) aims to develop a model of a realistic background error statistics for an atmospheric CO_2 data assimilation system. For atmospheric CO_2 data assimilation, the errors in the background are not only impacted by the uncertainties in the CO_2 transport but also by the spatial and temporal variability of the carbon exchange at the Earth surface. The background errors cannot be prescribed via traditional forecast-based methods used within the NWP community, as

these fail to account for the uncertainties in the carbon emissions and uptake, resulting in an overall underestimation of the errors.

As part of this objective a new approach is proposed whereby the differences between two CO_2 model concentrations are used as a proxy for the statistics of the background errors. These error statistics are subsequently used to assimilate observations from the Greenhouse gases Observing SATellite (GOSAT) into an operational atmospheric 4D-VAR system implemented at the European Centre for Medium Range Weather Forecasts (ECMWF). The ECMWF system is designed akin to a NWP-DA setup (*Engelen et al*. [2009]), in which CO_2 mixing ratios are assimilated along with relevant meteorological variables to obtain consistent estimates of 4-dimensional atmospheric CO_2 concentrations. The resultant 4D CO_2 fields are evaluated against independent observations of CO_2 from aircraft profiles and from the Total Column Carbon Observing Network (TCCON) to gauge the impact of the background error statistics on the 4D-VAR analysis.

In essence, this objective answers (Q1) from a true assimilation perspective and demonstrates the need to adapt existing DA methods to carbon science applications. More generally, the developed approach for parameterizing the background error statistics is also relevant for other trace gas assimilation applications, especially ones where the background errors are influenced by both atmospheric transport and the emission patterns.

Objective 2: Evaluate data assimilations systems against a batch setup for a CO_2 flux estimation problem

The second objective (Chapter 5, *Chatterjee and Michalak* [in prep.]) explores the impact of the numerical data assimilation framework on the accuracy and precision of inversion estimates. A comparison of DA approaches under the same inversion conditions, or knowledge of their advantages/disadvantages for the CO_2 flux estimation problem is lacking. Similarly, the applicability of DA and the equivalence of their estimates to those from batch inversion methods have not been demonstrated either.

Using a simple 1-dimensional advection-diffusion inverse problem, estimates from two advanced DA (an ensemble-filter and a variational) methods are compared and assessed against the estimates from a batch inverse modeling scheme. Experiments are specifically designed to identify the impact of the observations and/or the choices of the DA parameters (i.e., ensemble size, number of descent iterations) in order to isolate the degradation relative to the batch solution. No dynamical model is specified for the DA methods to keep the problem setup analogous to a real CO_2 flux estimation problem. Results of the experiments are used to discuss the advantages and disadvantages of the two DA methods relative to the batch as well as implications for solving a real CO_2 flux estimation problem.

This objective addresses (Q2) and highlights some of the critical aspects in the implementation of DA for the CO_2 flux estimation problem. More generally, the conclusions presented in this study are aimed to provide: (a) a guide to the CO_2 inverse

community, for deciding which DA approach to invest in, and (b) a platform for future DA inter-comparison efforts using real CO_2 datasets.

Objective 3: Development of a geostatistical ensemble square root filter for estimating fine scale CO_2 fluxes

The third objective (Chapter 6, *Chatterjee et al*. [in press]) is to adapt existing data assimilation methods to a geostatistical inverse modeling framework. This study introduces a Geostatistical Ensemble Square Root Filter (GEnSRF) as a prototypical ensemble filter designed for high-resolution flux estimation applications. Within the geostatistical inverse modeling (GIM; *Michalak et al*. [2004]) framework, the prior estimate of the flux from biospheric models and inventories is replaced by a linear combination of auxiliary variables related to CO_2 flux processes. The proposed filter, thus, eliminates reliance on a particular biospheric model or inventory by directly leveraging the information content from the observations.

GEnSRF is applied to a synthetic data study over North America, in which CO_2 surface fluxes are estimated at a high spatial ($1° \times 1°$) and temporal (3-hourly) resolution. Subsequently, the ensemble performance, both in terms of estimates and the estimation uncertainties, is benchmarked against a GIM setup. This allows isolation and quantification of the degradation in the estimates due to the numerical approximations and parameter choices in the ensemble filter. Advanced techniques, such as adaptive covariance inflation (*Anderson* [2009]) and localization (e.g. *Houtekamer and Mitchell* [2001]) are also implemented within GEnSRF to make it a state-of-the-art filter. Multiple sensitivity tests are conducted to identify the critical factors in stabilizing the ensemble

filter and develop guidelines for tuning the ensemble performance, specifically for CO_2 applications.

Specifically this objective focuses on (Q1) and (Q3) and develops a modified ensemble filter framework that can leverage the information content of large and complex data streams in the design of flux estimation inference problems. Additionally, this objective also contributes to (Q2) and reinforces the conclusions from Objective 2 as it provides insights into the behavior of an ensemble DA approach. More generally, the results point to key differences between the applicability of ensemble approaches to carbon cycle science relative to meteorological applications, where these tools were originally developed.

Objective 4: Implement the geostatistical ensemble square root filter for estimating surface fluxes using satellite measurements of CO_2 concentrations

The final objective (Chapter 7) assesses the utility of global CO_2 distributions from the Greenhouse gases Observing SATellite "IBUKI" (GOSAT; *Kuze et al*. [2009]; *Yokota et al*. [2009]) towards the estimation of fine-scale CO_2 surface fluxes. The GOSAT instrument is the first operational space-based instrument designed specifically for measuring the dry air mole fractions of CO_2 (XCO_2). Yet robust results with simulated GOSAT XCO_2 retrievals have been obtained only at large spatial scales of subcontinents or oceanic basins (e.g. *Maksyutov et al*. [2008], *Kadygrov et al*. [2009]; *Chevallier et al*. [2009b]; *Hungershoefer et al*. [2010]). As a consequence it is unclear: (a) whether GOSAT retrievals can constrain fine-scale fluxes with reasonable precision and accuracy, and (b) whether the dense but low precision GOSAT data really provides

additional information with respect to the high-precision but sparse observations from the surface monitoring network.

Observations retrieved from the GOSAT and CO_2 measurements from surface flask sites are assimilated using GEnSRF to estimate high-resolution (spatial: 1° × 1.25°; temporal: daily) global surface fluxes. GEnSRF is ideal for this application as it is capable of ingesting the high-density data and estimating grid-scale sources and sinks of CO_2 independent of any *a priori* flux estimates from biospheric model and inventories. The purely data-driven flux estimates provide unbiased quantification of the observational influence on the CO_2 flux estimates. Examining the analysis sensitivity with respect to the observations (*Cardinali et al.* [2004]) provides quantitative estimates of the constraints provided by the different observational datasets. Comparison of the resultant flux estimates with a bottom-up estimate is used to assess the source-sink information gleaned out from the GOSAT XCO_2 retrievals by GEnSRF.

While this objective primarily addresses (Q3) in providing flux estimates that eliminate reliance on a particular bottom-up model and/or inventory, it partially contributes to (Q1) in demonstrating a real data application of GEnSRF. The direct scientific impact of this objective, however, is not only the high-precision data driven flux estimates but also information about the constraints that GOSAT observations provide regarding global surface fluxes. As the GOSAT XCO_2 retrievals mature, the data-driven GEnSRF flux estimates is expected to provide valuable insights into the processes driving the net uptake and release of carbon in space and time.

The rest of the dissertation is organized in the following manner. Chapter 2 reviews the current state of carbon cycle science and the application of DA for atmospheric CO_2. Chapter 3 is intended to be a brief mathematical primer on inverse modeling and data assimilation approaches. Chapters 4 through 7 outline the specific studies framed around the four objectives discussed earlier. Chapter 8 concludes the dissertation and provides several research recommendations for further improving CO_2-DA applications.

Chapter 2

The Carbon Cycle –

Processes and Measurement Tools

2.1 INTRODUCTION

In the last century, humans have strived towards a better understanding of the existence of a "greenhouse" effect and the significance of the level of rising atmospheric CO_2 concentrations on the Earth's surface temperature. Although the potential alteration of the climate through anthropogenic activities was identified clearly by the middle of the twentieth century (e.g. *Fourier* [1824]; *Tyndall* [1861]; *Arrhenius* [1896]; *Chamberlin* [1897]; *Callendar* [1938]; *Revelle and Suess* [1957]), only over the last two decades have humans widely recognized the greenhouse effect and begun to respond collectively (*GCP* [2009]).

The IPCC Climate Change Synthesis Report (IPCC [2007]) identified CO_2 as the most important *anthropogenic* greenhouse gas. Since 1990, CO_2 alone has been responsible for approximately 80% of the change in the Earth's radiation balance amongst all major greenhouse gases (*Hofmann et al*. [2006]). Atmospheric CO_2 concentrations have increased from 280 ppm in pre-industrial times (circa 1750) to 392 ppm (~40% above pre-industrial) in 2011, with a mean growth rate of about 2.07 (± 0.09) ppm per year (http://www.esrl.noaa.gov/gmd/ccgg/trends/) within the last decade. Given the current level of increase, it is anticipated that a doubling of CO_2

from pre-industrial levels will occur by the middle of the century. This is expected to lead to an average global increase of temperature of 2.0-4.5 °C via a positive climate-carbon cycle feedback (e.g. *Friedlingstein and Prentice* [2010]; *Friedlingstein et al*. [2006]; *Cox et al*. [2000]). These predictions illustrate the importance, but do not necessarily show the multitude, of carbon-climate feedbacks possible (e.g. *Berthelot et al*. [2005]; *Matthews et al*. [2007]; *Heimann and Reichstein* [2008]; *Gloor et al*. [2010]; *Roy et al*. [2011]; *Bellard et al*. [2012]) or the uncertainty in the understanding of thresholds and tipping points in the Earth's carbon cycle (U.S. Carbon Cycle Science Plan, ed. *Michalak et al*. [2011]).

This uncertainty in the carbon-climate feedback primarily stems from a fundamental lack of understanding of the processes governing the carbon cycle. Current inaccurate assessments of the '*size, spatial distribution, uncertainty and likelihood of disturbance of carbon pools*' (*Canadell et al*. [2010]) also contributes to this uncertainty. Based on a survey of strategy reports aiming to synthesize and assess the role of the carbon cycle and its responses to human activities (e.g. U.S. Carbon Cycle Science Plan, ed. *Michalak et al*. [2011]; European Commission Report, ed. *Schulze et al*. [2009]; State of the Carbon Cycle Report, ed. *King et al*. [2007]), three key questions have been identified by the carbon science community, which need to be addressed:

C1. *What are the spatial and temporal patterns of carbon fluxes at large scales (continents, ocean basins)? How do regional and sub-regional patterns in carbon fluxes interact with the global-scale carbon cycle?*

C2. *What are the multiple mechanisms responsible for current aquatic and terrestrial carbon sinks? What are the relative contributions of these mechanisms, and their interactions?*

C3. *How will natural dynamics of the carbon cycle and human activities feedback to influence future atmospheric* CO_2 *concentrations?*

These science questions specifically aim to address the interactions between different carbon pools and the processes controlling their interactions. Process-level understanding will increase the ability to develop realistic scenarios of the carbon cycle that can be used to assess and anticipate future changes in carbon fluxes and atmospheric CO_2 concentrations. This, however, first requires an accurate diagnosis of global and regional carbon sources and sinks with realistic uncertainties.

Significant progress has already been made in identifying the major carbon pools and the flow between them at global scales. It is well-established that in the absence of anthropogenic emissions, the carbon exchange between the earth surface and the atmosphere are approximately balanced over multi-decadal time scales (*Denman et al*. [2007]). Both the land biosphere and the ocean continuously exchange CO_2 with the atmosphere through a variety of processes such as photosynthesis, respiration, air-sea gas transfer etc. Anthropogenic activities, especially fossil-fuel emissions and changes in land use, perturb the balance as these represent a net source of CO_2 to the atmosphere (Figure 2.1).

At global scales, the overall magnitude of these sources and sinks are relatively well-known. For example, the CO_2 community has reached an agreement that the Northern lands are responsible for the largest portion of the net terrestrial carbon sink (e.g. *Stephens et al*. [2007]). The strength of the net terrestrial sink is variable from one year to the next because the carbon balance of terrestrial ecosystems strongly responds to climate variability (e.g. *Bousquet et al*. [2000]; *Rödenbeck et al*. [2003a]; *Canadell et al*. [2004]) such as precipitation, surface temperature, radiation (e.g. *Le Quéré et al*. [2009]), and other inter-annual variability resulting from fires, land use changes etc. The average global terrestrial net uptake for the past three decades have been updated continuously as new methods and carbon cycle models have become

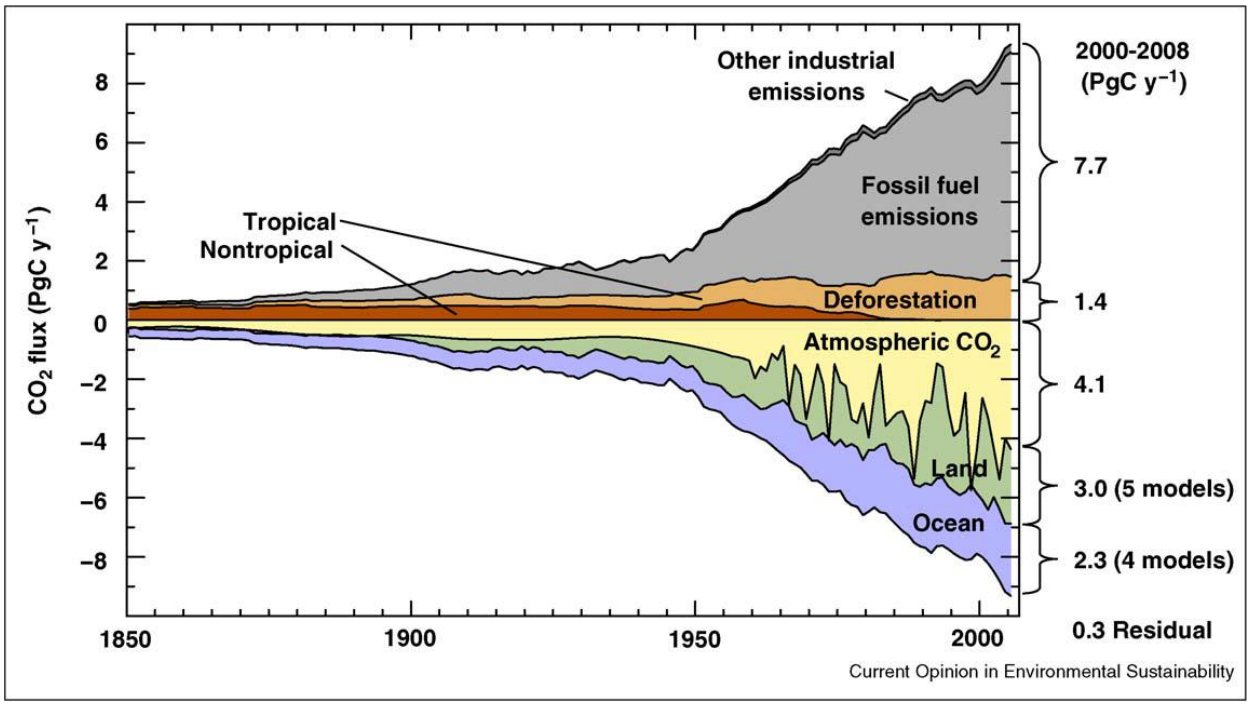


Figure 2.1*- This figure adapted from Raupach and Canadell [2010] show the major fluxes in the global carbon cycle over (roughly) the Anthropocene. The anthropogenic* CO_2 *emissions, comprising of fossil fuel emissions and deforestation, are shown as positive fluxes into the atmosphere. Conversely, the accumulation of carbon in three major components: the atmosphere, the terrestrial biosphere and the ocean are shown as negative fluxes. The numbers in the right panel give the 2000-2008 average estimates from Le Quéré et al. [2009].*

available (**1980s**: 0.2 PgC y^{-1}, *Houghton et al*. [2001]; ~0.27 PgC yr^{-1}, *Sarmiento et al*. [2010]; **1990s**: 1.2 PgC y^{-1}, *Houghton et al*. [2001]; 2.6±0.7 PgC yr^{-1}, *Canadell et al*. [2007]; 1.15 PgC yr^{-1}, *Sarmiento et al*. [2010]; **2008**: 4.7±1.2 Pg C yr^{-1}, *LeQuere et al*. [2009]) but overall remained consistent amongst different studies. Similarly in terms of the ocean budget, the agreement between different sets of estimates is closer than the land component. *Gruber et al*. [2009] provided two estimates of the net oceanic uptake of CO_2 for 1995 – 2002; the first based on climatological data (1.9 ± 0.6 PgC y^{-1}), and the second based on an inversion of interior ocean carbon observations (2.2 ± 0.3 PgC y^{-1}). *Manning and Keeling* [2006] used a novel approach, in which they examined the decline in atmospheric O_2/N_2 ratios and suggested an

oceanic uptake rate of 1.9 ± 0.6 PgC y^{-1} for 1990 to 2000 and 2.2 ± 0.6 PgC y^{-1} for 1993 to 2003, both of which are remarkably close to the estimates provided by *Gruber et al*. [2009].

Considerable uncertainties remain at regional scale budgets, however, especially over '*dynamic*' regions such as the tropics and the boreal zones on land and the high latitude and Southern Ocean regions (U. S. Carbon Cycle Science Plan, ed. *Michalak et al*. [2011]). Large uncertainties in the budget estimates at regional scales have hindered attempts to better understand the processes controlling the carbon cycle at local to regional scales and/or provide comprehensive verification of inventory estimates (e.g. *Canadell et al*. [2010]). Well-established regional budgets and their uncertainties not only help to attribute fluxes to underlying processes and drivers, but also to use the regional flux estimates to constrain the global budget. Reducing the uncertainty in regional budgets is also necessary for monitoring and verifying emissions at the scale of individual countries (e.g. *Dilling* [2007]; U. S. Carbon Cycle Science Plan, ed. *Michalak et al*. [2011]). Thus, a large amount of effort within the CO_2 research community is dedicated to better understanding the carbon cycle dynamics and their associated uncertainty at regional scales.

This dissertation is primarily driven by the need to identify CO_2 sources and sinks at regional (and/or finer) scales, as well as the need to further reduce uncertainties on existing carbon budget estimates. The methods developed as part of this dissertation are aimed at advancing specific tools used by the CO_2 research community to infer the exchange and magnitude of surface fluxes amongst different carbon pools. Against the backdrop of the key carbon science questions outlined earlier in (C1) to (C3), the contribution of this dissertation is an important cog in our efforts to reconcile the carbon sources and sinks with atmospheric CO_2 observations.

The rest of the chapter is organized as follows. Section 2.2 provides a high-level summary of the two primary methods ('bottom-up' and 'top-down') that are currently in vogue for obtaining carbon budgets. Section 2.3 reviews existing studies on the top-down (or atmospheric inverse modeling) approach, while Section 2.4 reviews the recent shift towards advanced data assimilation (DA) approaches for obtaining CO_2 flux estimates. Finally, Section 2.5 summarizes other DA applications for CO_2 science.

2.2 APPROACHES TO EVALUATE CARBON BUDGETS

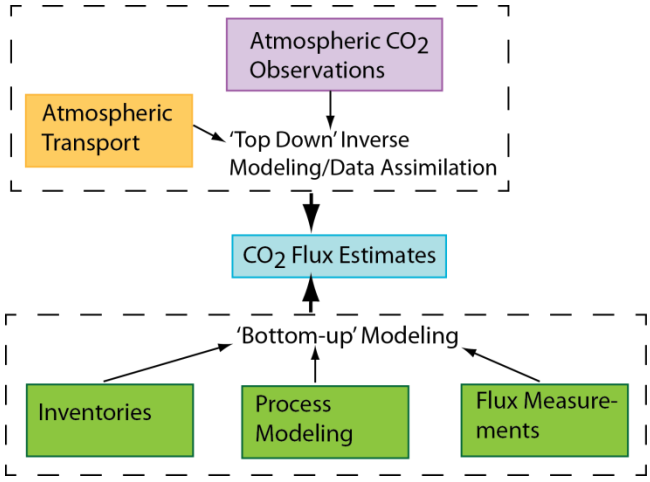


Figure 2.2*- Schematic showing the association between different techniques for obtaining estimates of surface carbon flux estimates.*

Knowledge of carbon fluxes and its magnitude can be inferred via two main approaches as shown in Figure 2.2 - "bottom-up" (based on measurements in the land or ocean compartments) and "top-down" (based on measurements in the atmosphere).

Bottom-up approaches are based on a process-based understanding of CO_2 surface exchange and can be broadly divided into three different methods - flux measurements, inventory datasets, and process models. The first method, direct measurement of carbon flux, is well developed over land for measurements over small spatial scales (for example, up to 1 km^2) using

the eddy correlation method (e.g. *Wofsy et al*. [1993]). Although eddy covariance measurements of carbon dioxide exchange between the biosphere and the atmosphere are now collected at more than 400 active sites (http://fluxnet.ornl.gov/site_status), the limited footprint of the flux tower network prevented accurate spatial scaling due to heterogeneity in land cover (e.g. *Raupach and Finnigan* [1995]; *Baldocchi* [2003]; *Friend et al*. [2007]; *Baldocchi* [2008]). Recent studies (e.g. *Jung et al*. [2011]; *Xiao et al*. [2012]), however, have demonstrated that novel machine learning techniques may be used to successfully upscale site-level data to global scales. Still, known issues with the eddy covariance technique, especially during night-time and under lateral flow conditions (e.g. *Saleska et al*. [2003]; *Alfieri et al*. [2011]) introduce large uncertainties in the measurements. Estimation of the magnitude of these uncertainties (e.g. *Finkelstein and Sims* [2001]; *Loescher et al*. [2006]; *Billesbach* [2011]), the variation in these uncertainties across different sites and regions (e.g. *Hollinger and Richardson* [2005]; *Richardson et al*. [2006]; *Luyssaert et al*. [2009]; *He et al*. [2010]), and/or their impact on inverse estimation of ecosystem model parameters (e.g. *Lasslop et al*. [2008]) remain active areas of research. Similar to land, over the oceans eddy flux measurements are possible (e.g. *Wanninkhof et al*. [2009]) but estimates based on measurements of surface water CO_2 partial pressure are more widely used (e.g. *Takahashi et al*. [2002]).

The second bottom-up method is inventory datasets, which are based on a variety of observations of specific flux components (i.e., gross productivity, respiration, land use change and disturbance, fossil fuel combustion, ocean fluxes etc.). The observations for each of the flux components range from measurement of trees on land (e.g. *Birdsey* [2004]; *McKinley et al*. [2011]) and land-use change (e.g. *Houghton* [2003]; *Houghton* [2010]), fossil fuel emissions from fuel sales, air pollution measurements (e.g. *Gurney et al*. [2009]) or night-time lights (e.g.

Oda and Makyutov [2011]), carbon in ocean-water samples (e.g. *Takahashi et al*. [2002]; *Sabine et al*. [2004]) etc. Even though inventory datasets can provide useful constraints on changes in the size of carbon pools, their utility is limited for quantifying trends longer than the last couple of decades. For example, the most complete forest inventories exist only since 1990 for northern mid-latitudes, where several plots have been systematically measured. Inventory data are available for tropical forests (e.g. *Houghton* [2005; 2007]; *Phillips et al*. [2009]) as well, though to a much lesser extent than for the northern temperate zones. Recent improvements in monitoring the above ground biomass via satellite observations (e.g. *Goetz et al*. [2009]), and model simulations (e.g. *Wulder et al*. [2008]) may allow more long-term estimates (e.g. *Pan et al*. [2011]), in the near future.

Finally, process-based models build from an understanding of the underlying processes of atmosphere/ocean or atmosphere/ecosystem carbon exchange to make estimates over scales of space and time that are relevant to the global carbon cycle. While initial studies (e.g. *Potter et al*. [1993]; *Denning et al*. [1996a]; *Ruimy et al*. [1996]) presented simple frameworks, these have been significantly refined and a variety of sophisticated models are now available, all of which differ substantially in the data used as constraints and/or the processes simulated (for example, see *Huntzinger et al*. [2012], *Keenan et al*. [2012] for discussions on a range of terrestrial biosphere models, and *Matsumoto et al*. [2004], *Gruber et al*. [2009] for discussions on a range of ocean models). The biggest advantage of the process-based models are that they allow for an easy understanding of the processes driving the carbon cycle (e.g. *Bondeau et al*. [2007]; *Piao et al*. [2009]; *Randerson et al*. [2009]). Yet the large variations across models estimates for both short and long-term projections (e.g. *Heimann et al*. [1998]; *Cramer et al*. [2001]; *Sitch et al*. [2008]; *Schwalm et al*. [2010]; *Dietze et al*. [2011]), imperfect parameterizations of the processes

represented in the models (e.g. *Braswell et al*. [2005]; *Xu et al*. [2006]; *Williams et al*. [2009]; *King et al*. [2011]; *Medvigy and Moorcroft* [2012]), and/or missing processes at large scales (e.g. *Moorcroft* [2006]; *Beer et al*. [2010]) cast doubt on the current capabilities of these models to provide robust regional budgets.

Juxtaposed to the bottom-up approaches, the basic premise of the top-down or inverse modeling approach is that - given a set of atmospheric CO_2 observations, and using a model of atmospheric transport, it is possible to infer information on the distribution of CO_2 sources and sinks at the surface of the Earth (e.g. *Enting* [2002]). In principle, given a perfect transport model together with an unlimited and perfectly distributed set of atmospheric CO_2 observations, the inverse modeling approach should provide CO_2 flux estimates at any desired spatial and temporal scale. But because both the transport model and the observational networks are imperfect, a multitude of different surface source/sink configurations become compatible with the atmospheric data. As a consequence, inverse modeling approaches for studying the carbon cycle vary significantly (e.g. *Ciais et al*. [2010b]) in the time scale of the analysis, the spatiotemporal resolution of the inferred fluxes, the inverse modeling framework used, and other setup choices. Since in this dissertation the main focus will be on the inverse modeling framework, the literature is reviewed in more detail in the next section (Section 2.3), while the mathematical framework is outlined in Chapter 3.

A long-term interest within the CO_2 community has been to reconcile the estimates from bottom-up and top-down approaches (e.g. *Canadell et al*. [2000]). Although the top-down approach ensures that the estimated fluxes are consistent with the atmospheric data, it does not take as much advantage of scientific understanding of plant physiology, ocean dynamics and other biogeochemical processes associated with the carbon cycle as is possible by the bottom-up

approaches. This may be disadvantageous, as it loses out on these sources of information and provides limited insights into the processes controlling the variability in the fluxes. The advantage, however, is that the top-down estimates have the potential to validate inventories (e.g. *Levin and Karstens* [2007]; *Desai et al*. [2011]; *Hayes et al*. [2012]), and/or identify errors in the process-based model formulations (e.g. *Peylin et al*. [2005]; *Desai et al*. [2010]). In recent years, a larger number of studies have aimed to reconcile the estimates from these two different approaches (e.g. *Baker et al*. [2008]; *Wang et al*. [2009]; *Ciais et al*. [2010a]; *Turner et al*. [2011]; *Gourdji et al*. [2012]), and coordinated regional synthesis activities are being pursued (for example, the North American Carbon Program - http://nacp.ornl.gov/int_synthesis.shtml, plus the REgional Carbon Cycle Assessment and Processes - http://www.globalcarbonproject.org/reccap/overview.htm). Initial results from these activities have already demonstrated the large spread that exists in the bottom-up and the top-down model estimates of CO_2 flux at both continental and sub-continental scales, due to differing model setups and input data. The soon to be completed RECCAP synthesis (*Canadell et al*. [2011]; *Patra et al*. [2012]) is expected to finally provide a better consistency check and provide new assessment of carbon fluxes and their drivers over different regions. As noted by *Canadell et al*. [2010], an ability to reconcile the CO_2 flux estimates from these two approaches would lead to not only minimizing the uncertainty within each approach but also increasing confidence in results, especially at intermediate regional scales (e.g. individual states, provinces, or countries). Providing independent sets of top-down flux estimates, however, is a significant challenge and will be an important motivating factor behind the methods developed and discussed later in this dissertation (Chapters 6 and 7).

2.3 'TOP-DOWN' APPROACHES TO STUDYING THE CARBON CYCLE

The approach most commonly adopted in atmospheric inverse modeling of CO_2 sources and sinks is based on Bayesian inverse modeling (see Chapter 3 - Section 3.3.1 for the mathematical framework). As outlined in the last section, the variations in the atmospheric CO_2 observations are traced back to the most likely configuration of source-sink patterns with the help of an atmospheric transport model.

Typically, atmospheric CO_2 observations are based on *in situ* measurements from a network of surface sites, ship observations and aircraft profiles, for example, those operated by the National Oceanic and Atmospheric Administration (NOAA – http://www.esrl.noaa.gov/gmd/dv/site/site_table2.html; *Tans and Conway* [2005]), or the Scripps Institute of Oceanography (*GLOBALVIEW*-CO_2 [2008]). Historically, CO_2 surface measurement sites have been deliberately placed in regions remote from terrestrial sources and sinks such as mountain tops and remote marine locations (e.g. *Fan et al*. [1998]). These sites are less exposed to anthropogenic and terrestrial influence and highly representative of the background CO_2 conditions. Continental sites do a better job of addressing the role of the local and regional influences on the atmospheric carbon cycle (e.g. *Gloor et al*. [2001]). Continental CO_2 records, especially from tall towers (e.g. *Bakwin et al*. [1998]) are available at hourly resolution and contain high-frequency information regarding the interaction between the terrestrial biosphere and the atmosphere at diurnal and synoptic (~days) time scales (e.g. *Geels et al*. [2004]). Although the surface measurements have high precision, they do have limited spatial coverage and are confined to the planetary boundary layer. Of late, the availability of aircraft measurements from targeted campaigns have allowed reasonable characterization of vertical gradients of atmospheric CO_2 (e.g. *Stephens et al*. [2007]; *Machida et al*. [2008]; *Niwa et al*.

[2011]; *Sawa et al*. [2012]). The vertical gradients are not only an indicator of the processes controlling the carbon exchange near the surface (i.e., under the planetary boundary layer) but also useful in evaluating model simulations of vertical and seasonal variations of atmospheric CO_2 (e.g. *Stephens et al*. [2007]; *Parazoo et al*. [2008]; *Miyazaki et al*. [2009]). Regular aircraft measurements, however, are limited to a few sites that remain unequally distributed around the globe (e.g. *Niwa et al*. [2012]).

The finite number of measurements available, coupled with atmospheric mixing, makes the CO_2 inverse problem both underdetermined (i.e., the total number of fluxes to be estimated is much greater than the number of observations available) and ill-posed (i.e., several different solutions exist that are equally consistent with the available measurements). To address the under-determined nature of the problem, historically, synthesis Bayesian inversions were constructed to estimate fluxes for a small number of pre-defined continental regions (e.g. *Bousquet et al*. [2000]; *Gurney et al*. [2003]; *Rödenbeck et al*. [2003b]; *Baker et al*. [2006b]; *Nassar et al*. [2011]). Specifying the spatial patterns of CO_2 flux within such large regions leads to aggregation errors (e.g. *Kaminski et al*. [2001]), whereby the atmospheric measurements are sensitive to variability in the fluxes at finer scales than the scale at which the inversion is allowed to adjust the fluxes. Inference of fluxes at large regions are also not amenable to capturing vegetation and climate variability at sub-regional scales, and thus provide little information on the mechanisms driving the underlying carbon cycle.

Newer studies have resolved this shortcoming by estimating fluxes at finer scales (i.e., 'biome-scale', 'grid-scale'). In some cases, like the ongoing Carbon Tracker project (*Peters et al*. [2007; 2010]), a set of scaling factors is used to adjust fluxes within a biome in order to match the observations, while in others, fine-scale (e.g. 8 km, 50 km etc.) fluxes have even been

estimated (e.g. *Gerbig et al*. [2006]; *Lauvaux et al*. [2009]; *Carouge et al*. [2010a]). Solving at such fine-scales, however, increases the computational challenge as a result of which the latter set of studies typically focus on specific regions within the continental United States or Europe. In between these two extremes lie the grid-scale inversions at 1° by 1°, 3.75° by 5°, 5° by 5°, or 8° by 10° etc. Often termed as 'regional scale' inversions, a suite of atmospheric inverse modeling studies using synthetic or real atmospheric CO_2 observations have been presented over the last decade (e.g., *Kaminski et al*. [1999]; *Gerbig et al*. [2003a; 2003b]; *Rödenbeck et al*. [2003a]; *Michalak et al.* [2004]; *Peylin et al*. [2005]; *Matross et al*. [2006]; *Mueller et al*. [2008]; *Gourdji et al*. [2010]; *Rivier et al*. [2010]; *Schuh et al*. [2010]; *Goeckede et al*. [2010]; *Gourdji et al*. [2012]), which have ranged from global to domain-specific application.

Alongside the under-determined nature of the problem, an important consideration is handling the ill-posed nature of the inverse problem. In order to extract a meaningful unique solution, additional information on the sources and sinks has to be introduced into the calculation. Examples of this additional information that are relevant to the CO_2 inversion problem include maps of air-sea fluxes from observations or ocean models (e.g. *Takahashi et al*. [2002]), patterns of terrestrial CO_2 exchanges inferred by terrestrial models (e.g. *Randerson et al*. [1997]; *Krinner et al*. [2005]), inventories of fossil fuel emissions (e.g. *Brenkert* [1998]; *Olivier and Berdowski* [2001]; *Andres et al*. [2012]), or combination of other bottom-up estimate from biospheric/oceanic models and inventories.

One of the reported issues with specifying *a priori* bottom-up flux estimates is that in the Bayesian framework, prior estimates will generally be recovered in areas under-sampled by the atmospheric network. As such, any assumptions associated with the biospheric models and inventories will be aliased onto the final flux estimates. Also, if the prior uncertainties are not

correctly specified, it can lead to spurious *a posteriori* flux estimates and/ or under-estimated *a posteriori* uncertainties (e.g. *Engelen et al*. [2002]; *Michalak et al*. [2005]).

These issues have prompted studies that relax some of the dependence on prior information from bottom-up estimates. The geostatistical inverse modeling approach (see Chapter 3 -Section 3.3.2 for the mathematical framework) to atmospheric CO_2 flux estimation (e.g. *Michalak et al*. [2004]; *Mueller et al*. [2008]; *Gourdji et al*. [2012]) is an example of one such approach. The underlying principle in the geostatistical approach is to specify a model of the trend, which can be a simple set of mean fluxes in space and time or consist of auxiliary environmental variables related to carbon flux. These auxiliary variables can range from vegetative indices such as Leaf Area index, Enhanced Vegetation index, Fraction of Photosynthetically Active Radiation, and/or indices associated with biomass burning, land use change and fossil fuel emissions (e.g. *Gourdji et al*. [2008; 2012]). The coefficients relating these auxiliary variables to the final fluxes are estimated as part of the inversion using the atmospheric data as a constraint. The geostatistical approach is thus more data-driven than its Bayesian counterpart and aims to leverage the information content of atmospheric observations more strongly.

Several other factors (for example - the atmospheric transport model, the choice of prior flux estimates for the Bayesian inverse modeling approaches, the parameterization of the covariance matrices, specification of boundary conditions for regional inversions etc.) are also known to control the quality of the flux estimates. An in-depth discussion on each of these topics is beyond the scope of this dissertation, and the reader is referred to *Ciais et al*. [2010b] for a more comprehensive summary on these issues. In spite of the different parameterization or setup choices amongst the different inversion studies, almost all of these have unanimously agreed that

the quality of the flux estimates is severely constrained by the sparse distribution of the *in situ* atmospheric CO_2 measurement network.

A promising opportunity for overcoming the sparse coverage of *in situ* measurements is satellite observations of atmospheric CO_2 concentrations. This option has been explored by the CO_2 community for over a decade now (e.g. *Rayner and O'Brien* [2001]). Missions currently operating and planned by international space agencies collect a broad range of data that aim to provide space-based measurements of CO_2 with dense coverage in space and time. These include passive sensors such as the Atmospheric Infrared Sounder (AIRS; *Chahine et al*. [2006]), infrared atmospheric sounding interferometer (IASI) on Met-Op-1 (*Crevoisier et al*. [2009]), scanning imaging absorption spectrometer for atmospheric cartography (SCIAMACHY; *Buchwitz et al*. [2005]), the greenhouse gases observing satellite (GOSAT; *Kuze et al*. [2009]), the Tropospheric Emission Spectrometer (TES; *Kulawik et al*. [2010]) and the planned re-flight of the Orbiting Carbon Observatory-2 (OCO-2; *Boesch et al.* [2011]; *Eldering et al*. [2012]). It is important to distinguish though between CO_2-dedicated missions such as GOSAT and OCO-2, and other multi-purpose missions such as TES, IASI and AIRS. The multi-purpose missions are typically more sensitive to the mid-troposphere, which makes it difficult to relate the measured variations in CO_2 concentrations to the spatial and temporal variations of CO_2 surface fluxes (e.g. *Chevallier et al*. [2009a]; *Ciais et al*. [2010b]). Future active sensors are also in the pipeline, such as the Active Sensing of CO_2 Emissions over Nights, Days, and Seasons (ASCENDS) satellite that is listed as a Tier 2 mission in the NRC Decadal Survey (*NRC* [2007]; *NASA* [2008]).

The current focus in the top-down community is to increase the resolution of flux estimates in both space and time to better understand the dynamics of the global carbon cycle.

This idea has been made possible by the increased availability of atmospheric CO_2 observations from recent remote-sensing instruments, and an expansion of the *in situ* network, over the last two to three years. Two limitations, however, have proved a bottleneck in making the most efficient use of the data. First the long life time of CO_2 makes its relative variations much smaller than for shorter-lived species. Such a small variability imposes a stringent constraint on the retrieval uncertainties from the remote-sensing instruments (e.g. *Rayner and O'Brien* [2001]; *Miller et al*. [2007]), which are yet to be attained (also see Chapter 7 for a more in-depth discussion on this aspect). Secondly, the, high-density data from the satellite necessitates usage of advanced data assimilation techniques in order to make the inverse problem computationally tractable. The development and application of data assimilation techniques for the carbon problem has spawned off its own area of research and studies, which are described in the next section.

2.4 DATA ASSIMILATION FOR CO_2 SOURCE-SINK ESTIMATION

The inverse modeling techniques mentioned above (i.e. Bayesian Inverse Modeling – e.g. *Enting* [2002] and Geostatistical Inverse Modeling – e.g. *Michalak et al*. [2004]) are employed in a batch mode, in which the inversion is performed in one step, by solving a single system of linear equations relating the CO_2 fluxes and the atmospheric CO_2 observations. The batch inverse modeling technique requires running an atmospheric transport model either once per estimated flux region/period combination or once per observation. This becomes computationally infeasible given the increasing spatial and temporal resolution of satellite data and ground based CO_2 concentration measurements, as well as the need for solving fluxes at fine spatiotemporal scales. The effort of pre-calculating and storing the transport model runs is becoming too large

even for today's supercomputers, and the resulting set of equations cannot be solved by traditional batch methods.

Computational challenges have created a need for moving away from batch setups to try and solve the system in a time stepping approach, where at every time step smaller subsets of unknown fluxes are optimized. This is exactly what has been pursued in the Numerical Weather Prediction (NWP) research for several decades now under the banner of data assimilation (DA). Although DA methods had made their way into the trace gas flux estimation problem in the early years of 2000 (e.g. *Kleiman and Prinn* [2000]; *Petron et al*. [2004]; *Yudin et al*. [2004]), application to the CO_2 flux estimation problem did not happen until 2005. Since then a series of studies have investigated methods based on the regular Kalman Filter (e.g. *Bruhwiler et al*. [2005]; *Michalak* [2008]; *Bruhwiler et al*. [2011]; *Zhuravlev et al*. [2011]), the ensemble Kalman Filter (e.g., *Peters et al*. [2005]; *Feng et al*. [2009]; *Kang et al*. [2011]; *Miyazaki et al*. [2011]; *Chatterjee et al*. [in press]), the variational scheme (e.g. *Chevallier et al*. [2005b]; *Rödenbeck et al*. [2005]; *Baker et al*. [2006a]) or hybrid approaches, such as the Maximum Likelihood Ensemble Filter (MLEF; *Zupanski et al*. [2007a]; *Lokupitiya et al*. [2008]).

It is important to point out though, that DA applications for CO_2 source-sink estimation are relatively new compared to batch inverse modeling applications. As a result, in most cases the capabilities of different DA systems and the skill of the DA inversions are still evolving. The majority of studies have been presented as proof-of-concept studies demonstrating the capabilities of individual DA systems using simulated data and/or estimating fluxes at coarse resolutions. These flux estimates have been subsequently evaluated by comparing them to biospheric model and inventory estimates and/or by assessing how well they reproduce available atmospheric CO_2 observations (e.g. *Peters et al*. [2005]; *Chevallier et al*. [2007]; *Lokupitiya et*

al. [2008]; *Feng et al*. [2009]; *Baker et al*. [2010]; *Kang et al*. [2011]; *Miyazaki et al*. [2011]). With the recent launch of the GOSAT instrument and availability of high-density data, it is expected that DA tools and their applications will evolve to answer more carbon science relevant questions.

From the viewpoint of this dissertation, the focus will be on ensemble Kalman filter (see Chapter 3- Section 3.4.1 for the mathematical framework) and the variational (see Chapter 3 – Section 3.4.2 for the mathematical framework) approaches, and hence applications for CO_2 source-sink estimation are presented in greater detail from the perspective of these two data assimilation approaches.

2.4.1 Source-sink studies using an ensemble Kalman filter

The most extensive application of the ensemble methods have been in the development of CarbonTracker (*Peters et al*. [2005]; http://www.esrl.noaa.gov/gmd/ccgg/carbontracker/) that has been used to reanalyze the recent flux history of CO_2 using an atmospheric transport model (TM5; *Krol et al*. [2005]) coupled to an ensemble square root filter (*Whitaker and Hamill* [2002]). Initially applied only for North America (*Peters et al*. [2007]), currently CarbonTracker has been extended to infer sources and sinks over Europe (*Peters et al*. [2010]) and is in the development stages for application to South America, Asia and Australasia (W. Peters, *pers. comm.*) and other atmospheric trace gas species (such as CH_4 - L. Bruhwiler, *pers. comm.*). CarbonTracker starts by forecasting atmospheric CO_2 mole fractions around the globe from a combination of CO_2 surface exchange models and the TM5 transport model driven by meteorological fields from the European Centre for Medium-Range Weather Forecasts (ECMWF). The resulting three-dimensional CO_2 distribution is then sampled at the time and location when *in situ* observations are available, and the difference between observations and

model forecast is minimized. This minimization is achieved by tuning a set of linear scaling factors that control the magnitude of the surface fluxes for larger, but sub-continental eco-regions based on the classification proposed by *Olson et al*. [2001]. Once the value of each of the scaling factors is determined in many consecutive assimilation cycles, the history of surface CO_2 exchange at 1° × 1° is reconstructed. Similar to CarbonTracker applications, *Feng et al*. [2009] have used a stochastic ensemble filter to estimate 8-day CO_2 surface fluxes over 144 geographical regions using simulated XCO_2 measurements.

More recent studies (e.g., *Miyazaki et al*. [2011]; *Kang et al*. [2011]; *Chatterjee et al*. [in press]) have used variants of the ensemble Kalman Filter but with various advancements in the filter algorithm and/or their assimilation characteristics. Contrary to CarbonTracker though, almost all of these newer studies have solved for fluxes at high resolutions but using only simulated measurements. For, example both *Miyazaki et al*. [2011] and *Kang et al*. [2011] solve for weekly fluxes at the model resolution (~2.8°) using the Local Ensemble Transform Kalman Filter (LETKF; *Hunt et al*. [2007]). *Chatterjee et al*. [in press] solve for 3-hourly fluxes at a 1° × 1° spatial resolution using a geostatistical variant of the ensemble square root filter (EnSRF – *Whitaker and Hamill* [2002]) that is designed to minimize the influence of information from biospheric models and/or inventories. The development of the geostatistical filter is a key component of this dissertation and will be discussed in greater detail in Chapter 6 (with a synthetic data application over North America) and Chapter 7 (with a real-data application over the globe assimilating satellite retrievals).

2.4.2 Source-sink studies using a variational scheme

The introduction of variational schemes to the carbon flux estimation problem (e.g. *Chevallier et al*. [2005b]; *Rödenbeck et al*. [2005]; *Baker et al*. [2006a]) happened around the

same time as ensemble methods. Yet in terms of their data usage, the 4D-VAR applications have been more varied than their ensemble counterparts. Almost all types of CO_2 observations, including *in situ* (e.g. *Rödenbeck et al*. [2005]; *Chevallier et al*. [2010a]), simulated satellite (e.g. *Chevallier et al*. [2007]; *Baker et al*. [2006a; 2010]), actual satellite observations (e.g. *Chevallier et al*. [2005a; 2009a]), or retrievals of the CO_2 total column from the Total Column Carbon Observing Network (TCCON; e.g. *Chevallier et al*. [2011]) have been assimilated.

The 4D-VAR setup in these studies is similar in most respects to the basic formalism used in the NWP community, except that instead of optimizing an initial condition (i.e., the atmospheric state) at the start of a relatively short assimilation window, in most cases, time-varying boundary values (surface CO_2 fluxes) are optimized over a longer span. The use of a long time window is necessary to account for the long mixing time scales of the CO_2 fluxes. Furthermore, this is possible due to the linearity of the CO_2 flux estimation problem unlike the non-linear evolution of the atmospheric state for the NWP problem.

Unlike the ensemble schemes, which are being developed worldwide by different groups, the development and application of 4D-VAR has been consigned to a few specific groups. The drawback to a wider implementation of 4D-VAR within the CO_2 inverse community has been the need for the development and maintenance of an adjoint of the atmospheric transport model. Of the available 15-25 [1] CO_2 transport models and variants thereof, to the best knowledge of the author only four global transport models - LMDz, PCTM, TM5 and GEOS-Chem (*Chevallier et al*. [2005b]; *Baker et al*. [2006a]; *Meirink et al*. [2008]; *Henze et al*. [2007]) have corresponding adjoints, and only two of these (LMDz and GEOS-Chem) are up-to-date with their respective

[1] *This number is approximate based on a review of published studies over the last five years, e.g. Baker et al. [2006a], Geels et al. [2007], Sarrat et al. [2007], Houweling et al. [2010]; Law et al. [2010], Niwa et al. [2011], Saito et al. [2011] etc. that have inter-compared global and regional transport models for CO_2. In reality, there may be a larger number of regional-scale transport models that have been developed worldwide.*

forward models. Difficulties associated with the availability of the adjoint stem largely from a general lack of efforts within the CO_2 community in developing/maintaining atmospheric transport models (e.g. *Rayner* [2010]). In this dissertation, even though a geostatistical variant of the 4DVAR technique was developed, it could not be applied to any real data problem because of the lack of an adjoint.

Finally, studies that compare different DA methods (ensemble-filter and 4D-VAR) under the same conditions are lacking. Hence, the advantages/disadvantages of different DA techniques with respect to solving the CO_2 flux estimation problem still remains to be evaluated (e.g. *Rayner* [2010]; *Zupanski et al*. [2007a]). This existing gap is filled in by Chapter 5 of this dissertation (*Chatterjee and Michalak* [in prep.]) where two state-of-the-art DA methods are implemented and compared for a source-sink estimation problem. By identifying the advantages and disadvantages of the two DA methods, it is expected that the findings will provide valuable guidance to the CO_2 research community, in deciding which approach to invest in, for future source-sink estimation purposes.

Given the increasing coverage and resolution of the remote-sensing data, DA provides the most computationally efficient scheme to characterize CO_2 sources and sinks on regional scales. The key challenge though will be in: (a) delivering the satellite data with high enough precision ($\sim \leq 1.5$ ppm) to yield new constraints on CO_2 fluxes, and (2) having advanced DA tools to optimally use this data and quantify regional sources and sinks with sufficient accuracy to reduce the uncertainty of current carbon budgets (or, in other words address carbon science question, C1). As stated earlier, reducing this uncertainty will enable us in a better understanding of the process-based impacts (C2), improving the mechanistic models for future predictions of carbon-climate feedbacks and implement effective mitigation policies (C3).

2.5 OTHER APPLICATION OF DATA ASSIMILATION IN CO_2 SCIENCE

Earlier in Chapter 1, it was mentioned that aside from using DA for its computational efficiency in the design of flux estimation inference problems, a few studies apply DA methods to estimate atmospheric CO_2 concentrations or parameters of biogeochemical models. These studies are amongst the first applications in CO_2 science, and any discussion on CO_2 -DA would be incomplete without highlighting these briefly.

The first of these, which has been pursued at the European Centre for Medium-Range Weather Forecasts (ECMWF) is to obtain consistent estimates of atmospheric CO_2 concentrations (e.g. *Engelen et al*. [2004; 2009]), using 4D-VAR to assimilate CO_2 observations along with a host of meteorological variables. The setup of this system is akin to a NWP-like system and poses analogous challenges regarding error and bias characterization etc. In fact Chapter 4 (*Chatterjee et al*. [in prep.]) in this dissertation explores this system in detail, and proposes a new approach for parameterizing the background error statistics of this system. The real benefit of this system lies in generating global CO_2 fields, which enhances the observational database, as the DA procedure uses physical and dynamical laws, along with the available observations, to constrain the analysis. The global CO_2 fields are generated with high-precision and these can be subsequently used to constrain CO_2 surface fluxes via inverse modeling (e.g. *Chevallier et al*. [2009a]).

The second of these has been the construction and application of a terrestrial Carbon Cycle Data Assimilation (http://ccdas.org/). The original framework proposed by *Kaminski et al*. [2002] has been expanded over the last decade (e.g. *Rayner et al*. [2005]; *Scholze et al*. [2007]; *Knorr et al*. [2010]; *Ziehn et al*. [2011b]) but its core concept of using atmospheric observations to constrain the parameters of a terrestrial biosphere model, has remained unchanged. Via a 4D-

VAR setup, CCDAS essentially combines the top-down and bottom-up approaches in calibrating key model parameters that: 1) improves the underlying behavior and the predictive capability of the model, and 2) can be used to provide estimates of high-resolution surface CO_2 fluxes, if necessary. The same conceptual framework has been extended to evaluate the optimality of CO_2 networks (*Kaminski et al*. [2012]) as well as specifically estimate photosynthetic parameters using CO_2 and heat fluxes observed locally at eddy-covariance sites (*Santaren et al*. [2007]).

Chapter 3

Mathematical Concepts

3.1 INTRODUCTION

Data Assimilation (DA) is essentially a class of numerical methods that has been widely used by the geosciences community to estimate the state of a system, given a discrete model of the dynamics of the system, a (noisy) estimate of the current state and (noisy) observations of the system over time. The data assimilation problem can also be considered as a generalized inverse problem[1] (e.g. v*an Leeuwen and Evensen* [1996]), although typically the inverse modeling paradigm follows a slightly different representation. As per *Rodgers* [2000], "*the inverse problem is the question of finding the best representation of the required parameter given the measurements made, together with any appropriate prior information that may be available about the system and the measuring device*". Thus, the emphasis is more on retrieving system parameters or time-varying boundary values of the system, rather than the initial state or systems' forecast.

In this chapter, DA principles are introduced from a generic inverse modeling standpoint, primarily because much of the conceptual implementation in this dissertation is based on the

[1] *Proponents of non-linear data assimilation have stated that data assimilation is not an inverse problem (e.g. van Leeuwen [2010]). Within the geosciences community, the terms data assimilation and inverse problem have been used interchangeably due to their formulation based on a least squares framework (e.g. Lewis et al. [2006]). But the way the observations and errors are modeled and handled for these two frameworks, and/or the way the prior or background information is brought into play has resulted in subtle distinctions between these two frameworks (e.g. Bocquet [2005]). Also see footnote [2] in Section 3.2.*

atmospheric inverse modeling paradigm. Section 3.2 defines a generic inverse problem, and develops the objective functions required to obtain a solution to the inverse problem. Solutions are discussed in terms of two batch inverse modeling schemes in Section 3.3, and two data assimilation (sequential and variational) schemes in Section 3.4. In this chapter only the general methodological aspects of the different inverse modeling/data assimilation schemes are described, with occasional references to CO_2 applications. A standard set of notations (Appendix A) are introduced for all the equations, which will be followed in this chapter and throughout the dissertation as well.

3.2 FORMAL STATEMENT OF THE PROBLEM

Assume one has a set of *m* parameters, represented by a state vector **s** ($\mathbf{s} \in \mathbb{R}^m$) that one wishes to infer from a set of *n* observations, represented by a vector **z** ($\mathbf{z} \in \mathbb{R}^n$). The two vectors are related by a forward model *h*, which may be an atmospheric transport model, an observation operator, a radiative transfer model etc., or a combination thereof, that encapsulates our understanding of the physics of the system:

$$\mathbf{z} = h(\mathbf{s}) + \boldsymbol{\varepsilon}_r \qquad 3.1$$

where $\boldsymbol{\varepsilon}_r$ accounts for the possible mismatch between the observations and the model's prediction of the observations. The generic inverse problem is to estimate the parameter vector **s**, knowing the observation vector **z** and its error characteristics.

In the majority of real-world applications, however, the number of observational data is lower than the number of parameters to be estimated ($n << m$) that makes the problem under-determined. Also the inverse problem is ill-posed whereby several different solutions and not one unique solution exist, which are equally consistent with the available measurement data (e.g.

Tarantola [2005]). A common approach is to regularize the solution by providing prior [2] information about the parameters, where the *a priori* expected value is represented as $E[\mathbf{s}] = \mathbf{s}^b$ ($\mathbf{s}^b \in \mathbb{R}^m$). Like all other information sources, the prior (or the background) has some error associated with it and this is captured through the prior (or background) error statistics $\boldsymbol{\varepsilon}_q$. Thus, the generic inverse problem is redefined as finding the set of parameter values that optimally balances the information from the observational data and the prior (or the background) along with their associated errors.

3.2.1 Bayesian objective function

Using a Bayesian framework, all the sources of information that can be used to constrain the estimate **s** are represented in terms of probability density functions (*pdf*'s). Three key assumptions will be made now: (a) the forward model *h* is linear, (b) both the measurement (or the likelihood) *pdf* $P(\mathbf{z} \mid \mathbf{s})$ and the prior *pdf* $P(\mathbf{s})$ can be captured by Gaussian statistics, and (c) the error statistics of the observation and the prior (or background) are centered ($E[\varepsilon_r] = E[\varepsilon_q] = 0$), mutually uncorrelated ($E[\varepsilon_r, \varepsilon_q] = 0$), and have respective variances $E[\varepsilon_r^2] = \mathbf{R}$ and $E[\varepsilon_q^2] = \mathbf{Q}^b$.

Note that in general, Bayesian estimation theory does not require specific assumptions regarding the form of the *pdf's*. Rather it only provides a framework to formalize the conditional relationship of the *pdf's* and conjoin the two sources of information from the observation and the

[2] *The terms 'prior' and 'background' are often used interchangeably in the inverse modeling community as they represent a subset of all possible parameter states that are more likely due to some pre-existing knowledge of the physical process being estimated. But in operational data assimilation, 'prior' refers to a set of initial conditions specified at the start of the assimilation (very first time step). For subsequent time steps, the term 'background' is used and refers to a model forecast of the analysis from previous time steps. Specifically, for this chapter I will use the term 'prior (or background)' where these can be used interchangeably. Else if the terms prior (i.e. initial conditions at the very first time step) and background (i.e. model forecast) is used separately, then they do have separate meanings.*

prior (or background). Assumption of Gaussian *pdf's* is common for modeling probability distributions as it is both algebraically convenient and a reasonable approximation for the real errors (e.g. *Rodgers* [2000]). Similarly, assumptions of a linear *h* are valid for the atmospheric CO_2 inverse problem, in which case it is typically represented by its linearized matrix form **H**. For a larger set of inverse problems encountered in geophysical applications *h* may not be linear and is only assumed in the vicinity of the background state, i.e., $h(\mathbf{s}) - h(\mathbf{s}^b) \approx \mathrm{h}(\mathbf{s}\text{-}\mathbf{s}^b)$.

Using the above assumptions, the likelihood *pdf* and the prior *pdf* are typically written as:

$$P(\mathbf{z}\,|\,\mathbf{s}) = a_1 \exp\left(-\left(\mathbf{z}-h(\mathbf{s})\right)^T \mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right)/2\right) \qquad 3.2$$

$$P(\mathbf{s}) = a_2 \exp\left(-\left(\mathbf{s}-\mathbf{s}^b\right)^T \left(\mathbf{Q}^b\right)^{-1}\left(\mathbf{s}-\mathbf{s}^b\right)/2\right) \qquad 3.3$$

where **R** and $\mathbf{Q}^b$ are the $n \times n$ model-data mismatch covariance and the $m \times m$ prior (or background) error covariance. Usually the model-data mismatch covariance matrix **R** is structured as a diagonal matrix and aims to capture the combined measurement, forward model, representation and aggregation errors (e.g. *Engelen et al*. [2002]). The measurement error is the actual error associated with the measuring instrument, and for remote-sensing observations, can also include errors from the retrieval algorithm. The model error is intended to capture the error in *h* either due to wrongly specifying the forward model parameters or due to an oversimplification of the real physics and dynamics in the forward model. The representation error describes the error accrued due to the fact that the actual observations are representative of a smaller scale relative to the resolution of the forward model. Finally, the aggregation error component is the error accrued due to estimating parameters at a coarser resolution, in both space and time. Low (or high) values along the diagonal of this matrix thus determine whether the inversion must reproduce the observations to a high (or low) precision. The $\mathbf{Q}^b$ matrix represents

the uncertainties associated with the prior (or background) estimates of the parameters and can have pre-defined spatial and temporal correlations relating the different parameters of the state vector to each other.

Using Bayes' theorem, the aim is to calculate the *a posteriori pdf* $P(\mathbf{s}\,|\,\mathbf{z})$ by updating the prior *pdf* with the measurement *pdf*. Since the prior (or the background) and the observational fields are two independent sources, the most likely estimate of the state is given by the joint probability of the above two probabilities:

$$P(\mathbf{s}\,|\,\mathbf{z}) = \frac{P(\mathbf{z}\,|\,\mathbf{s})P(\mathbf{s})}{\int P(\mathbf{z}\,|\,\mathbf{s})P(\mathbf{s})\,d\mathbf{s}} = a_3 \exp\left(-J(\mathbf{s})\right) \tag{3.4}$$

Combining Equations 3.2 to 3.4, $J(\mathbf{s})$ is defined as:

$$J(\mathbf{s}) = \frac{1}{2}\left(\mathbf{z}-h(\mathbf{s})\right)^T \mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right) + \frac{1}{2}\left(\mathbf{s}-\mathbf{s}^b\right)^T \left(\mathbf{Q}^b\right)^{-1}\left(\mathbf{s}-\mathbf{s}^b\right) \tag{3.5}$$

Due to our assumption of Gaussian *pdf*'s, maximizing the joint probability in Equation 3.4 is the same as minimizing the *objective function* (or the *cost*, or *penalty function*) in Equation 3.5. Essentially higher values of $J(\mathbf{s})$ correspond to less probable states. The contributions of the observational term denoted by J_o and the prior (or the background) term, denoted by J_b is distinguished as:

$$J(\mathbf{s}) = \underbrace{\frac{1}{2}\left(\mathbf{z}-h(\mathbf{s})\right)^T \mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right)}_{J_o} + \underbrace{\frac{1}{2}\left(\mathbf{s}-\mathbf{s}^b\right)^T \left(\mathbf{Q}^b\right)^{-1}\left(\mathbf{s}-\mathbf{s}^b\right)}_{J_b} \tag{3.6}$$

The Bayesian objective function (Equation 3.5 or 3.6), thus, represents the sum of least-squared difference between the parameters to be estimated and the prior (or the background) and observational fields, weighed by their respective error covariance matrices.

3.2.2 Geostatistical objective function

A slightly different objective function (e.g. *Kitanidis and Vomvoris* [1983]; *Michalak et al*. [2004]) can be obtained by re-defining the way the prior (or the background) information is represented in the system. This definition follows from the field of geostatistics where the parameter **s** is modeled as a random vector with *a priori* expected value (i.e., mean):

$$E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta} \tag{3.7}$$

X is a known $m \times p$ matrix representing different processes, which influence the parameters to be estimated, $\boldsymbol{\beta}$ are $p \times 1$ unknown drift coefficients, and $\mathbf{X}\boldsymbol{\beta}$ is the model of the mean. The prior *pdf* is now modeled as:

$$P(\mathbf{s},\boldsymbol{\beta}) = a_2 \exp\left(-(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})^T \left(\mathbf{Q}^b\right)^{-1} (\mathbf{s}-\mathbf{X}\boldsymbol{\beta})/2\right) \tag{3.8}$$

In Equation 3.8, note that: a) the drift parameters $\boldsymbol{\beta}$ are assumed to be unknown as well and hence these need to be estimated along with the original set of parameters **s**, and b) $\mathbf{Q}^b$ now describes the expected variability in the departures of the parameters from the model of the trend. Other than the changes in the prior *pdf*, the likelihood function (Equation 3.2) and the Bayesian scheme (Equation 3.4) remains the same and the geostatistical objective function is defined as:

$$J(\mathbf{s},\boldsymbol{\beta}) = \frac{1}{2}\left(\mathbf{z}-h(\mathbf{s})\right)^T \mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right) + \frac{1}{2}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})^T \left(\mathbf{Q}^b\right)^{-1}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}) \tag{3.9}$$

Analogous to the Bayesian objective function, this Geostatistical objective function (Equation 3.9) also represents a compromise between reproducing the observations and staying close to a statistical model of the trend ($\mathbf{X}\boldsymbol{\beta}$), where the covariance matrices determine the relative weight of these competing objectives.

The objective functions defined in Equations 3.6 and 3.9, however, fail to single out one particular estimate for **s**. In order to select a particular solution of **s**, it is necessary to specify an estimator. Possible examples are the class of approaches that are based on finding the "minimum

variance estimate" or based on the "maximum *a posteriori* (MAP) estimate". The minimum variance estimators minimize the variance of the analysis error, thus ensuring that the minimum is the most probable state. Conversely, the MAP solution is based on finding the state that maximizes the posterior *pdf* $\hat{\mathbf{s}} = \arg\max P(\mathbf{s}|\mathbf{z})$. For a linear problem with Gaussian statistics, these two frameworks lead to two equivalent algorithms because of the symmetrical properties of the *pdf*. Since the numerical implementations of the estimators are very different, this leads to different solutions schemes, as discussed in the next two sections.

As an aside, using a subset of applications from the literature, Box 1 defines the various ingredients of the objective functions (i.e., observations, prior or background information, error covariance matrices, forward model etc.) in terms of an atmospheric CO_2 inverse modeling application or an atmospheric CO_2 data assimilation application.

3.3 BATCH INVERSE MODELING SCHEMES

The batch inverse modeling schemes are based on finding the minimum variance estimate and aim to solve the system of (linear) equations in a single step. Two distinct sets of equations, based on whether the Bayesian or the Geostatistical objective function is used, can be identified.

3.3.1 Bayesian inverse modeling (BIM)

The minimum of the Bayesian objective function defined in Equation 3.6 can be obtained by taking the derivative with respect to the parameters **s**, and setting them to zero. The final solution (e.g. *Enting* [2002]) in the form of *a posteriori* means and covariances is given by:

BOX 1 – Ingredients of the objective function based on a subset of the CO_2 literature

Variable (Dimensions)	Atmospheric CO_2 Inverse Modeling Application [*]	Atmospheric CO_2 Data Assimilation Application [**]
s ($m \times 1$)	Discretized surface flux distribution, in space and time	4D atmospheric fields, including CO_2 concentrations
z ($n \times 1$)	CO_2 observations from *in situ* and/or satellite	Observed radiances or CO_2 observations from *in situ* and/or satellite; meteorological observations
h	Linear and a combination of the transport model and the observation operator. Its matrix form **H** ($n \times m$) is also commonly used.	*Non-linear* and a combination of the radiative transfer model and an observation operator
R ($n \times n$)	Model-data mismatch covariance	Model-data mismatch covariance
$\mathbf{Q}^b$ ($m \times m$)	Covariance of flux deviations from either the prior estimate $\mathbf{s}^b$ or the model of the mean $\mathbf{X\beta}$	Covariance of the background error due to our imperfect knowledge of the 'true' errors in the atmospheric state
$\mathbf{s}^b$($m \times 1$)	Prior (or background) estimate from a process based model output (i.e., biospheric fluxes) or inventory data (i.e., ocean air-sea CO_2 exchange, fossil fuel emissions etc.)	*Background state* usually taken from a short range forecast valid at the initial time
X($m \times p$**)**	Auxiliary variables such as vegetative indices (land cover, LAI, fPAR etc.) or socioeconomic datasets (population density, fossil fuel emissions, GDP etc.). Simpler - unknown but constant mean, i.e., $\mathbf{X} = [1 \ldots . 1]^T$	n/a

(* e.g. *Michalak et al*. [2004], *Ciais et al*. [2010b] ** *Engelen et al*. [2009])

$$\hat{\mathbf{s}} = \mathbf{s}^b + \mathbf{K}\left(\mathbf{z} - \mathbf{H}\mathbf{s}^b\right) \quad 3.10$$

$$\mathbf{Q}^a = \left(\mathbf{I} - \mathbf{K}\mathbf{H}\right)\mathbf{Q}^b \quad 3.11$$

$$\mathbf{K} = \mathbf{Q}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{Q}^b\mathbf{H}^T + \mathbf{R}\right)^{-1} \quad 3.12$$

where **H** is the $n \times m$ linearized matrix form of h that represents the sensitivity of the observations **z** to the parameters **s** (i.e., $H_{i,j} = \partial z_i / \partial s_j$), **K** is the Kalman gain, $\hat{\mathbf{s}}$ is the posterior best estimate of the parameters, $\mathbf{Q}^a$ is the *a posteriori* covariance of the recovered best estimate, with the diagonal elements of $\mathbf{Q}^a$ representing the predicted error variance ($\sigma^2_{\hat{s}}$) of individual elements in $\hat{\mathbf{s}}$. This form of the estimate is commonly termed as the *best linear unbiased estimator* (BLUE) due to the fact that the term **K** that weights the observation increments, can be analytically shown to be the optimal gain (*Jazwinski* [1970]). The dimension of the matrix to be inverted in Equation 3.12 is $n \times n$. If the number of parameters to be estimated is significantly larger than the number of observations, it is computationally efficient to use this form. Alternate formulations (e.g. *Enting* [2002]) of Equations 3.10 - 3.12 exist, in which the dimension of the matrix to be inverted is $m \times m$; this formulation is computationally cheaper if the number of parameters is significantly smaller than the number of observations.

3.3.2 Geostatistical inverse modeling (GIM)

As per Equation 3.9, the inverse problem has two vectors of unknowns. To find the minimum variance estimator, the geostatistical objective function is minimized with respect to the state and the drift parameters, **s** and **β**, respectively. Algebraic manipulation (e.g. *Michalak et al*. [2004]) yields the following system of linear equations:

$$\begin{bmatrix} \mathbf{H}\mathbf{Q}^b\mathbf{H}^T + \mathbf{R} & (\mathbf{H}\mathbf{X}) \\ (\mathbf{H}\mathbf{X})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} (\mathbf{H}\mathbf{Q}^b) \\ \mathbf{X}^T \end{bmatrix} \quad 3.13$$

Once the above system is solved for $\mathbf{\Lambda}$ ($m \times n$) and $\mathbf{M}$ ($p \times m$), they are used to obtain the *a posteriori* estimate of the parameters $\hat{\mathbf{s}}$ and their posterior covariance $\mathbf{Q}^a$:

$$\hat{\mathbf{s}} = \mathbf{\Lambda}\mathbf{z} \quad 3.14$$

$$\mathbf{Q}^a = \mathbf{Q}^b - \mathbf{Q}^b\mathbf{H}^T\mathbf{\Lambda}^{\mathrm{T}} - \mathbf{X}\mathbf{M} \quad 3.15$$

Note that in these equations, $\mathbf{M}$ is a matrix of Lagrange multipliers and should not be confused with *M*, which will be used to define the dynamical model later (Equation 3.19-3.20). Finally, the drift coefficients $\hat{\boldsymbol{\beta}}$ ($p \times 1$) and their uncertainties are estimated as:

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\left(\mathbf{Q}^b\right)^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\left(\mathbf{Q}^b\right)^{-1}\mathbf{\Lambda}\mathbf{z} \quad 3.16$$

$$\mathbf{Q}^a_{\hat{\boldsymbol{\beta}}} = (\mathbf{X}^T\mathbf{\Psi}^{-1}\mathbf{X})^{-1} \quad 3.17$$

where the diagonal elements of $\mathbf{Q}^a_{\hat{\boldsymbol{\beta}}}$ represent the uncertainty of the individual drift coefficients, and the off-diagonal terms represent the estimated covariance of the errors associated with these estimates.

The estimate can also be expressed in a fashion similar to the Bayesian best estimate. In this case, the best estimate is the sum of a deterministic component (i.e., model of the mean ($\mathbf{X}\hat{\boldsymbol{\beta}}$)) and a stochastic component that is a function of the *a priori* correlation structure in $\mathbf{Q}^b$:

$$\hat{\mathbf{s}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Q}^b\mathbf{H}^T\mathbf{\Psi}^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{X}\hat{\boldsymbol{\beta}}\right) \quad 3.18$$

Associated with the GIM framework, an additional task is the selection of auxiliary variables to be included in $\mathbf{X}$. These variables can be selected based on scientific understanding of the problem or one can use advanced variable selection methods such as hypothesis-based

techniques (e.g. Variance-Ratio Test -*Kitanidis* [1997]; *Gourdji et al*. [2008]) or criterion-based selection methods (e.g. Akaike information criterion - *Akaike* [1974], Bayes information criterion - *Schwarz* [1978]; *Mueller et al*. [2010]; *Yadav et al*. [2010a]).

Finally, in concluding this section it is worth highlighting that the matrix form of *h* (i.e., **H**) is necessary to solve both the BIM and the GIM systems. For the atmospheric CO_2 inverse problem, this is obtained by running an atmospheric transport model either once per estimated flux region/period combination, or once per observation if an adjoint to the transport model is available. Pre-calculating and storing this sensitivity matrix, however, is not a trivial task. For example, if one aims to estimate daily global fluxes at 1° × 1° for a month (360*180*30 = ~2e6) using atmospheric CO_2 observations from a satellite (~5000 observations per month), then the dimensions of **H** will be approximately [5e3 × 2e6]. The corresponding dimensions of $\mathbf{Q}^b$ will be [2e6 × 2e6] while the dimensions of **R** will be [5e3 × 5e3]. The cost of generating **H** and the sheer size of the matrices and their inverses involved makes the batch inverse modeling schemes computationally infeasible for such large-scale applications, necessitating the use of data assimilation schemes.

3.4 DATA ASSIMILATION SCHEMES

Given the vastness (and richness!) of DA methods and concepts, this section can by no means be regarded as a complete overview of the subject. Along with the specific references outlined in each section, the reader is encouraged to look at the excellent reviews of *Ghil and Malanotte-Rizzoli* [1991], *Cohn* [1997], *Todling* [1999] and specific textbooks by *Daley* [1991], *Rodgers* [2000], *Lewis et al*. [2006], *Lahoz et al*. eds. [2010] for a thorough introduction to the field.

For the purposes of this dissertation, a suite of DA schemes have been examined (Figure 3.1). It will not be possible to outline the mathematical framework of all the examined methods. Specifically among the Monte-Carlo based DA approaches, the Kalman filter is first introduced in order to motivate the ensemble square root filter (EnSRF). Amongst the variational DA approaches, both the 3-dimensional primal (3D-VAR) and dual (3D-PSAS) forms are discussed followed by the 4-dimensional variational method (4D-VAR). Note that all these methods will be discussed from the perspective of solving the objective function outlined in Equation 3.6 (i.e. regular Bayesian objective function) but can be theoretically converted to solve Equation 3.9 (i.e. Geostatistical objective function). The geostatistical approach adds another set of *p* parameters that need to be estimated along with the original state parameters. Given that the dimension of *p* is significantly smaller than *m*, the additional computational cost is trivial in most cases. The specific development of a geostatistical ensemble square root filter (GEnSRF) is presented in Chapter 6.

3.4.1 Sequential data assimilation scheme

The Kalman filter (*Kalman* [1960]) is an optimal data assimilation method for problems with linear dynamics and is derived based on the minimum variance requirement. The regular Kalman filter equations can be divided into two components: time update and measurement update equations.

Time update equations from time step $t_{\tau-1}$ to t_{τ}:

$$\mathbf{s}_{t_\tau}^{f} = M\hat{\mathbf{s}}_{t_{\tau-1}}^{a} \tag{3.19}$$

$$\mathbf{Q}_{t_\tau}^{f} = M\mathbf{Q}_{t_{\tau-1}}^{a}M^{T} + \left(\boldsymbol{\varepsilon}_{M}\right) \tag{3.20}$$

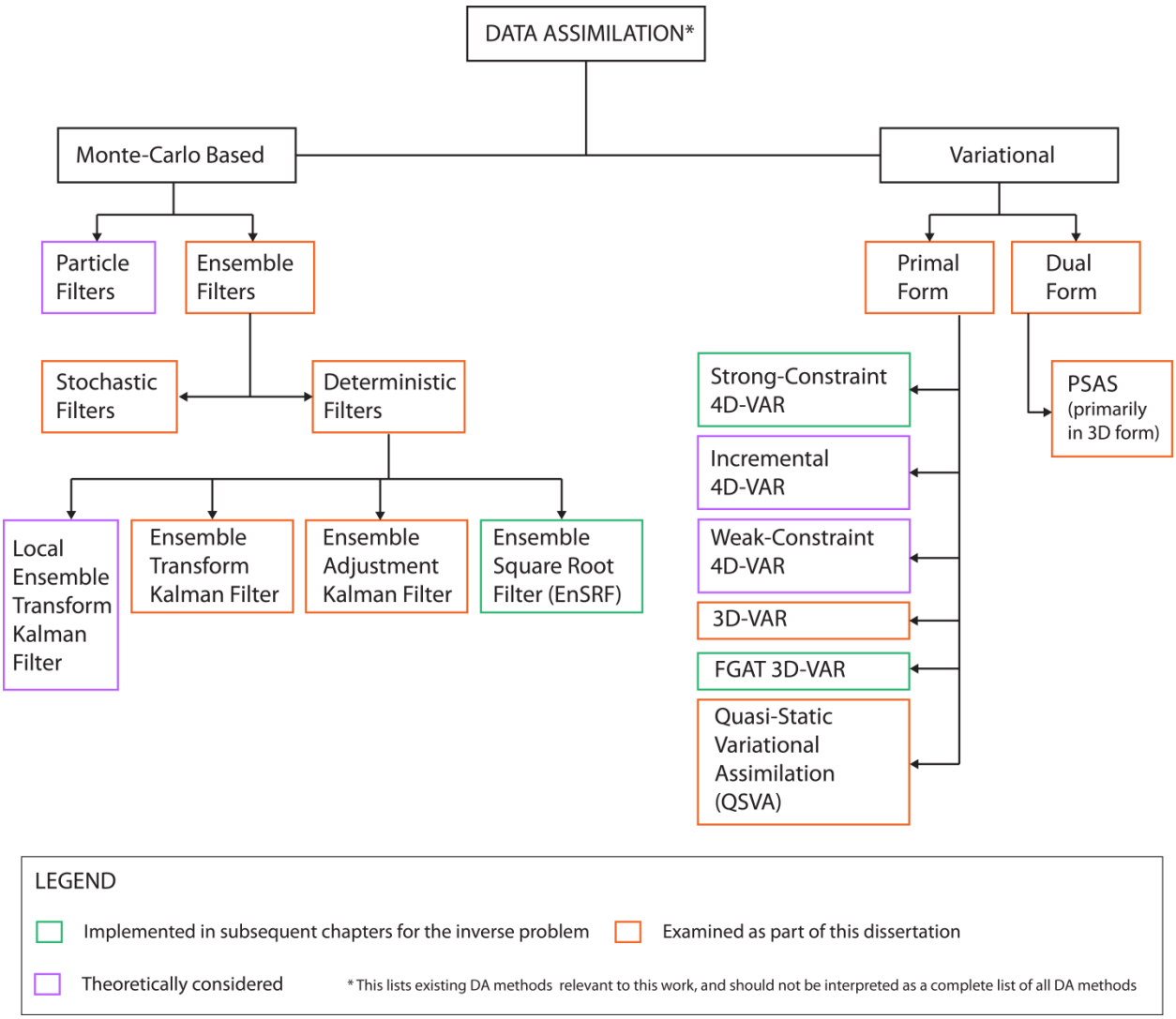


Figure 3.1 *– A subset of data assimilation algorithms that were examined in the dissertation. Specifically, in this chapter the mathematical formulation of the EnSRF (among the Monte-Carlo based methods) and the 3D-VAR, PSAS and strong constraint 4D-VAR (among the variational based methods) will be discussed.*

Measurement update equations at time step t_τ:

$$\mathbf{K}_{t_\tau} = \mathbf{Q}^f_{t_\tau}\mathbf{H}^T_{t_\tau}(\mathbf{H}_{t_\tau}\mathbf{Q}^f_{t_\tau}\mathbf{H}^T_{t_\tau} + \mathbf{R}_{t_\tau})^{-1} \tag{3.21}$$

$$\hat{\mathbf{s}}^a_{t_\tau} = \mathbf{s}^f_{t_\tau} + \mathbf{K}_{t_\tau}(\mathbf{z}_{t_\tau} - \mathbf{H}_{t_\tau}\mathbf{s}^f_{t_\tau}) \tag{3.22}$$

$$\mathbf{Q}^a_{t_\tau} = \left(\mathbf{I} - \mathbf{K}_{t_\tau}\mathbf{H}_{t_\tau}\right)\mathbf{Q}^f_{t_\tau} \tag{3.23}$$

M [3] is a dynamical/forecast model that propagates the parameters in **s** between two time steps, $\boldsymbol{\varepsilon}_M$ is the dynamical model error covariance, **I** is the identity matrix, $(\bullet)^f$ represents the forecast values at time t_τ and $(\bullet)^a$ represents the analyzed values at time t_τ.

A close examination of the measurement update equations indicates their equivalence to the BIM setup presented in Equations 3.10-3.12. At a particular time step, if the time indices are dropped, the Kalman gain matrix **K** is exactly similar to the term multiplying $\left(\mathbf{z}-\mathbf{H}\mathbf{s}^b\right)$ on the RHS in Equation 3.10. For the Kalman filter, the forecast $\mathbf{s}_{t_\tau}^f$ is the analyzed values from the previous time step propagated through the dynamical model to the current time step. At a particular time step t_τ, this is equivalent to the prior (or background), $\mathbf{s_p}$ in the Bayesian IM setup. It is not surprising then that the posterior covariance calculation of the Kalman filter is almost as difficult to implement as the Bayesian IM equations. Two main challenges can be identified: (i) the computational cost associated with the multiplication of the large matrices involved, and (ii) the need to calculate and pre-store the transport model runs **H**. Thus, practical implementation for a fine-scale CO_2 flux estimation problem is still not feasible with the Kalman filter scheme.

The ensemble Kalman filter (EnKF; *Evensen* [1994]) is a reduced-rank approximation to the optimal, full-rank Kalman filter solution that addresses these problems by using ensemble representations for the background and posterior error covariances. Ensemble size limits the

[3] *For the atmospheric CO_2 inverse problem, no suitable dynamical model is available; hence estimates of future CO_2 fluxes do not normally depend on the analysis of current CO_2 fluxes (Peters et al. [2005]). If one assumes $M = \mathbf{I}$, then the time update in Equations 3.19 and 3.20 can be ignored and one is concerned only with the measurement update equations in Equations 3.21-3.23. The absence of a dynamical model is not applicable for the CO_2 assimilation problem, where a full non-linear forecast model is available (e.g. Engelen et al. [2009]) to propagate the atmospheric state, including the CO_2 concentrations, between different time steps.*

number of degrees of freedom used to represent these error covariances and the calculations can be shown to be computationally practical for modest-sized ensembles (e.g. *Anderson* [2009]). Several flavors of EnKF are in vogue but with assumptions of linearity and Gaussian statistics, all of these flavors are closely related (*Anderson* [2003]). Specifically for a linear problem, deterministic filters have been shown to be more accurate than their stochastic counterparts (e.g. *Whitaker and Hamill* [2002]; *Lawson and Hansen* [2004]). *Tippett et al*. [2003] conducted a review of three popular deterministic filters (ensemble transform Kalman filter – ETKF, *Bishop et al*. [2001]; ensemble adjustment Kalman filter – EAKF, *Anderson* [2001]; ensemble square root filter – EnSRF, *Whitaker and Hamill* [2002]) and concluded that beginning with the same prior (or background) error covariance these different deterministic filters ultimately produce analyses ensembles that span the same state subspace and have the same covariance. The only difference is that the computational cost of ETKF and EAKF are dependent on the rank of the prior (or background) covariance matrix while the computational cost of EnSRF is dependent on the number of observations. This is highly attractive if the number of observations is lower than the number of parameters to be estimated, which makes EnSRF more computationally efficient relative to other deterministic filters.

Similar to the Kalman filter, the EnSRF algorithm derives the analysis equations by requiring that the total analysis error variances are minimized. Assuming the prior (or background) covariance is known, one can create *an ensemble of N state fields* $\mathbf{s}'^{b}$ (where N<<m) that have a background covariance $\mathbf{Q}^{b}$. These deviations $\mathbf{s}'^{b}$ are created such that the normalized ensemble of deviations defines the columns of the matrix $\mathbf{\Gamma}$ [$m \times N$] that is the square root of the covariance matrix $\mathbf{Q}^{b}$**:** $\mathbf{Q}^{b} = \mathbf{\Gamma}\mathbf{\Gamma}^{T}$

$$\mathbf{\Gamma} = \frac{1}{\sqrt{N-1}}\left(\mathbf{s}_1'^b, \mathbf{s}_2'^b \cdots \mathbf{s}_N'^b\right) \tag{3.24}$$

The matrix square roots of $\mathbf{Q}^b$ are not unique, and can be computed in different ways, for example through Cholesky decomposition or singular value decomposition. A convenient representation is to write:

$$\mathbf{Q}^b \approx \frac{1}{N-1}\left(\mathbf{s}'^b\right)\left(\mathbf{s}'^b\right)^T \tag{3.25}$$

In the limit of $N \to \infty$ this representation of $\mathbf{Q}^b$ is exact. For the measurement update step at time t_τ, the state vector and the deviations from the state vector are updated as:

$$\hat{\mathbf{s}}^a = \mathbf{s}^f + \mathbf{K}\left(\mathbf{z} - h\left(\mathbf{s}^f\right)\right) \tag{3.26}$$

$$\hat{\mathbf{s}}'^a = \hat{\mathbf{s}}'^f - \tilde{\mathbf{K}}\left(h\left(\mathbf{s}'^f\right)\right) \tag{3.27}$$

where the time indices have been dropped. Comparing Equation 3.26 to Equation 3.22, note that the forward model *h* is used directly, instead of using its matrix form **H.** Here **K** is the traditional Kalman gain matrix (similar to Equation 3.12) and $\tilde{\mathbf{K}}$ is a reduced form of **K**:

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{Q}^b\mathbf{H}^T + \mathbf{R}}}\right)^{-1} \mathbf{K} \tag{3.28}$$

The Kalman gain matrix **K** is used to update the state vector while the reduced form $\tilde{\mathbf{K}}$ is used to update the ensemble. Since the ensemble is a reduced rank representation of the full covariance, if the Kalman gain matrix were to be used to update the ensemble, it would result in a systematic underestimation of the covariances (e.g. *Whitaker and Hamill* [2002]). For the calculation of the gain matrices $\mathbf{Q}^b\mathbf{H}^T$ and $\mathbf{H}\mathbf{Q}^b\mathbf{H}^T$ is approximated by directly running the forward model (*h*) and sampling it with the ensemble (i.e. $h(\mathbf{s}'^b)$). If observations are being assimilated sequentially, $\mathbf{H}\mathbf{Q}^b\mathbf{H}^\mathbf{T}$ is a scalar value while the size of $\mathbf{Q}^b\mathbf{H}^\mathbf{T}$ is $m \times 1$:

$$\mathbf{HQ}^b\mathbf{H}^{\mathbf{T}} \approx \frac{1}{N-1}\left(h\left(\mathbf{s}'^b\right)\right)\left(h\left(\mathbf{s}'^b\right)\right)^T \tag{3.29}$$

$$\mathbf{Q}^b\mathbf{H}^{\mathbf{T}} \approx \frac{1}{N-1}\left(\mathbf{s}'^b\right)\left(h\left(\mathbf{s}'^b\right)\right)^T \tag{3.30}$$

In essence the innovations in Equations 3.29 and 3.30 ultimately reduce the computational cost of the ensemble scheme relative to the regular Kalman filter.

Once all observations have been assimilated, the *a posteriori* covariance $\mathbf{Q}^a$ is reconstructed from the ensemble (Equation 3.31). This explicit calculation of *a posteriori* error covariances is often advertised as one of the primary advantages of the ensemble-based data assimilation schemes, relative to variational schemes.

$$\mathbf{Q}^a = \frac{1}{N-1}\left(\mathbf{s}'^a\right)\left(\mathbf{s}'^a\right)^T \tag{3.31}$$

Once the measurement update step is completed, the time update steps as presented in Equations 3.19-3.20 are applied, and the ensemble filter is ready for the next measurement update step.

The ensemble adjustment Kalman filter (EAKF) of *Anderson* [2001] and the ensemble transform Kalman filter (ETKF) of *Bishop et al*. [2001], also uses the traditional Kalman filter update equation to update the ensemble mean (Equation 3.26). In the EAKF, a linear operator $\mathbf{A} = \mathbf{I} - \tilde{\mathbf{K}}\mathbf{H}$ is applied to the prior (or background) ensemble to get an updated ensemble whose sample covariance will be identical to the posterior covariance obtained via Equation 3.23. For the ETKF, a more complicated transformation matrix is used to update the ensemble. Several choices of the transformation matrix have been proposed in the literature (e.g. *Bishop et al*. [2001]; *Wang and Bishop* [2003]; *Wang et al*. [2007]) but the idea in all of them is to obtain an analysis error covariance, which is consistent with the full-rank Kalman filter analysis error covariance given by Equation 3.23. Thus all of the deterministic filters are functionally

equivalent but differ only in algorithmic details (e.g. *Whitaker and Hamill* [2002]; *Tippett et al*. [2003]).

An important feature of all ensemble filters is their dependency on the ensemble size. Given that the number of ensemble members N is finite, the representation of $\mathbf{Q}^b$ in N-dimensional space is not perfect. This results in a varying number of off-diagonal correlations (termed sampling error) as a consequence of which, a state variable may be incorrectly impacted by an observation that is physically remote. Several techniques have been proposed to account for the spurious noise in the ensemble (e.g. *Houtekamer and Mitchell* [2001]; *Hamill et al.* [2001]; *Ott et al*. [2004]; *Anderson* [2007b]). The most common approach is to perform a Schur (element wise) multiplication of a compactly supported correlation matrix ρ (*Gneiting* [2002]; *Bergemann and Reich* [2010]) with the covariance model generated by the ensemble:

$$\mathbf{Q}^b\mathbf{H}^T \approx \rho \bullet \left[\frac{1}{N-1} \left(\mathbf{s}'^b\right) \left(h\left(\mathbf{s}'^b\right)\right)^T \right] \qquad 3.32$$

The choice of the correlation matrix ρ can range from simple exponential based function to a more complex fifth-order Gaspari-Cohn function (*Gaspari and Cohn* [1999]). Generally, compactly supported functions are preferable as these functions are non-zero in only a small (local) region specified by a length scale thus ensuring that correctly specified physical correlations are maintained, while spurious correlations beyond the length scale are removed. More recently sophisticated balance-aware localization schemes have been proposed primarily for NWP applications (e.g. *Bishop and Hodyss* [2011]; *Kepert* [2011]; *Jun et al*. [2011]) as well as adaptive sampling error correction techniques based on an examination of the prior ensemble (e.g. *Anderson* [2012]).

Along with sampling error, limited ensemble size may lead to an ensemble collapse and filter divergence. Ensemble collapse (e.g. *Lawson and Hansen* [2004]; *Leewunburgh et al*.

[2005]; *Sacher and Bartello* [2008]) refers to the fact that before the assimilation the spread of the ensemble members reflects the prior (or background) error structure but after assimilation they tend to converge to a single value. The precise cause of ensemble collapse is not well understood but it is easy to see its manifestation through the Kalman filter equations. Specifically, equations 3.21 to 3.23 imply that as $\mathbf{Q}_{t_1}^f \to 0$, then $\hat{\mathbf{s}}_{t_t}^a \to \hat{\mathbf{s}}_{t_t}^f$; that is, if the prior (or background) uncertainty is very small, then the prior (or background) ensemble effectively informs the Kalman filter that one is "*very confident*" in the prior (or background), in which case the observations add very little information about the parameters to be estimated and the assimilation reduces to the prior (or background).

Several ad-hoc and adaptive techniques have been proposed in different ensemble-related applications to counter this loss of variance (e.g. *Anderson and Anderson* [1999]; *Zhang et al.* [2004]; *Hamill and Whitaker* [2005]; *Anderson* [2007a]; *Li et al*. [2009]; *Anderson* [2009]; *Pena et al*. [2010]; *Miyoshi* [2011]). In all ensemble-filter applications used in this dissertation, the adaptive technique proposed by *Anderson* [2009] will be used (also see Chapter 6 – Section 6.2.4).

3.4.2 Variational data assimilation scheme

3.4.2.1 3D-VAR and 3D-PSAS

Unlike sequential data assimilation schemes, the principle of variational data assimilation (e.g. *Sasaki* [1970 a, b and c]; *Lorenc* [1986]; *Talagrand and Courtier* [1987]) is based on rewriting the least squares problem posed in Equation 3.6 as a minimization problem that can be solved (approximately) using explicit minimization algorithms. This is also equivalent to finding the maximum *a posteriori* (MAP) estimate of the objective function. The variational method

works in an iterative fashion, whereby an estimate of the parameters in time are first run through the forward model to derive modeled measurements. These modeled measurements are then compared to the true measurements and the measurement residuals (weighted using the measurement error statistics) are subsequently run backwards in time through the adjoint model to obtain the adjustments to be made to the estimate of the parameters. The minimization can be stopped by limiting artificially the number of iterations or by requiring that the norm of the gradient decreases by a predefined amount during the minimization. The simplest implementation of this class of algorithms is in terms of solving the objective function in Equation 3.6. Termed as 3D-VAR (3-dimensional variational), in this technique both the prior (or background) $\mathbf{s}^b$ and the observational set $\mathbf{z}$ are assumed to be obtained at the same time. This is a reasonable assumption if either the time period over which the observations are available is short enough, or the underlying process can be considered stationary over this period (e.g. *Daley and Barker* [2001]; *Talagrand* [2010]). Minimization of the objective function requires the availability of the gradient of J with respect to the parameters $\mathbf{s}$:

$$\nabla J(\mathbf{s}) = \left(\mathbf{Q}^b\right)^{-1}\left(\mathbf{s}-\mathbf{s}^b\right) + \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right) \tag{3.33}$$

A particularly efficient way of calculating this derivative is through the adjoint approach (*Errico* [1997]), in which explicit calculation of all the elements of $\mathbf{H}^T$ is avoided by running the forward model backwards to compute the relationship between a small change in J and changes in $\mathbf{s}$. A single pass of the adjoint of the forward model computes $\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{z}-h(\mathbf{s})\right)$, which is then added to $\left(\mathbf{Q}^b\right)^{-1}\left(\mathbf{s}-\mathbf{s}^b\right)$ in Equation 3.33 above.

The second derivative of the objective function or the Hessian is given as:

$$\nabla^2 J(\mathbf{s}) = \left(\mathbf{Q}^b\right)^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \tag{3.34}$$

Calculating the Hessian is not strictly necessary for solving the variational problem. Most state-of-the-art minimization algorithms (e.g. L-BFGS, *Nocedal and Wright* [2006]) tend to approximate the Hessian to avoid the large-scale computational costs associated with the calculation and storage of the second derivatives. Recovering the exact Hessian is attractive though as its inverse corresponds to the analysis error covariance matrix $\mathbf{Q}^a$ (e.g. *Le Dimet et al*. [2002]; *Rödenbeck* [2005]), which is highly desirable for inverse problems. In most applications of variational algorithms either the uncertainty estimates are not reported or a suite of sensitivity tests are done to provide an approximate range within which the estimates may lie. Recent studies (e.g. *Shutyaev et al*. [2009]; *Cheng et al.* [2010]) have shown that computationally efficient alternatives for calculating the Hessian may exist, although these are yet to be applied to an inverse estimation problem.

A second type of variational algorithm can be derived based on Equation 3.10, which is re-written as:

$$\mathbf{s}-\mathbf{s}^b=\mathbf{Q}^b\mathbf{H}^T\left(\mathbf{HQH}^T+\mathbf{R}\right)^{-1}\left(\mathbf{z}-h\left(\mathbf{s}^b\right)\right) \tag{3.35}$$

Equation 3.35 is split into the following two equalities:

$$\mathbf{s}-\mathbf{s}^b=\mathbf{Q}^b\mathbf{H}^T\mathbf{w}, \textit{ where } \quad \mathbf{w}=\left(\mathbf{HQ}^b\mathbf{H}^T+\mathbf{R}\right)^{-1}\left(\mathbf{z}-h\left(\mathbf{s}^b\right)\right) \tag{3.36}$$

where $\mathbf{w}$ has the same dimension as $\mathbf{z}$ and is regarded as an increment in observation space while $\mathbf{Q}^b\mathbf{H}^T$ acts a smoothing term that maps the increment from the observation to the parameter space. Equation 3.36 allows us to define a new objective function on the dual of the observation space:

$$J(\mathbf{w})=\frac{1}{2}\mathbf{w}^T\left(\mathbf{HQ}^b\mathbf{H}^T+\mathbf{R}\right)\mathbf{w}-\mathbf{w}^T\left(\mathbf{z}-h\left(\mathbf{s}^b\right)\right) \tag{3.37}$$

This is known as the Physical Space Assimilation System (PSAS; e.g. *Da Silva and Guo* [1996]; *Cohn et al*. [1998]), which was first developed and implemented at NASA's Data Assimilation Office (DAO) in the late 1990s. Since the size of the control vector of the 3D-PSAS functional is determined by the number of observations *n* instead of the dimension of the parameters *m*, this makes it especially attractive if the number of observations is significantly lower than the number of parameters.

The dual approach (3D-PSAS) has been less popular than the primal approach (3D-VAR) even though studies (e.g. *Courtier* [1997]; *El Akkraoui et al*. [2008]) have pointed out that both formulations have the same convergence properties. A possible cause may be that the dual approach requires the use of $\mathbf{H}^T$ in Equation 3.36, which in the case of variational assimilation, implies strict linearity of the forward model operator. Conversely an argument in favor of the dual approach is that the inverses of the error covariance matrices (i.e., $\mathbf{R}$ and $\mathbf{Q}^b$) are not required, unlike 3D-VAR. As a result PSAS has no singularities when the error covariance matrices tend to zero (e.g. *Auroux* [2007]). This is important, especially when the 3D-PSAS is expanded to the 4D-PSAS, and the dynamical model error comes into play. The same PSAS algorithm can be used for both strong- (assuming no dynamical model error) and weak-constraint (assuming dynamical model error) variational assimilation. Recently, a secondary algorithm called the Restricted Preconditioned Conjugate Gradient (RPCG) has been proposed by *Gratton and Tshimanga* [2009], which presents an alternate framework to solve the linear variational problem in the dual space. Clearly, depending on the conditions of the problem and especially on the relative dimensions of the state and observation space, it may be more advantageous to use the dual form over the primal form, for many large-scale inverse problems.

3.4.2.2 4D-VAR

The 4 dimensional-variational or 4D-VAR in short, is the current state-of-the art variational approach where the parameters are estimated by minimizing the misfit between a temporal sequence of states and observations that are available over the given assimilation window (Figure 3.2). The four dimensional nature of 4D-VAR reflects the fact that the observation set spans not only three dimensional state spaces but also a time domain. Mathematically, this calls for a generalization of the 3D-VAR objective function (based on Equation 3.6) in time:

$$J(\mathbf{s}) = \sum_{t=t_0}^{t=t_\tau} \left(\mathbf{z}_t - h_t(\mathbf{s}_t)\right)^T \mathbf{R}_t^{-1} \left(\mathbf{z}_t - h_t(\mathbf{s}_t)\right) + \left(\mathbf{s} - \mathbf{s}^b\right)^T \left(\mathbf{Q}^b\right)^{-1} \left(\mathbf{s} - \mathbf{s}^b\right) \quad 3.38$$

The observations are now distributed among *t* times in the interval but the prior (or background) $\mathbf{s}^b$ and the analysis are defined only at the initial time. The parameters $\mathbf{s}_t$ at any given time may be identified as:

$$\forall t, \quad \mathbf{s}_t = M_{0\to t}(\mathbf{s}) \quad 3.39$$

Similar to Equation 3.19, the role of *M* is to forecast the parameters from the initial time to *t*. Commonly known as the *strong-constraint* 4D-VAR, this forms the basis for the variational assimilation systems widely employed within the numerical weather prediction (NWP) community (e.g. *Rabier* [2005]). In case the error associated with *M* cannot be neglected, a different objective function is necessary that adds an additional term (e.g. *Tremolet* [2006; 2007]). This approach is then known as the *weak constraint* 4D-VAR.

In the absence of the model forecast operator, a common assumption that can be made is $M = \mathbf{I}$. The 4D-VAR application is then known as the *First-Guess-At-the-right-Time* 3D-VAR, or FGAT3D-VAR. This process is four-dimensional in the sense that the observations distributed

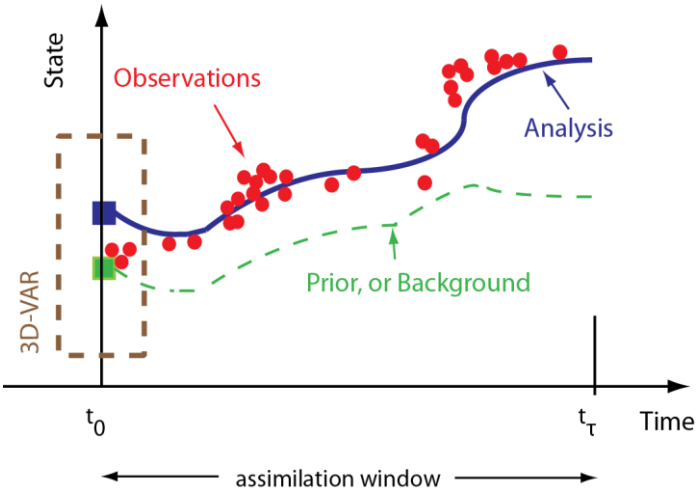


Figure 3.2 *- Schematic representation of a variational data assimilation scheme. Observations are collected over an assimilation window during which a prior, or background from the previous forecast, is also available. Observations update the prior (or background) to obtain a new analysis valid at the initial time. Difference between 3D-VAR and 4D-VAR is simply based on whether all observations over the assimilation window are assumed to be at the same time (3D-VAR), or assimilated at their correct times within the assimilation window (4D-VAR).This figure has been modified from Bouttier and Courtier [1999].*

over the assimilation window are compared with their analogues (i.e. modeled observations) at the correct time. But it is three-dimensional in that the minimization of the objective function does not use any explicit dynamics. Compared to 3D-VAR, the FGAT 3D-VAR still shows an improvement in the estimation quality (*Massart et al*. [2010]) because the temporal sequence of innovation vector it uses is more exact than the standard 3D-VAR.

In the context of 4D-VAR, there are many innovative ideas that have been explored and developed over the last couple of decades, especially within the NWP community. Operational systems using 4D-VAR for weather data assimilation have employed a number of algorithms (nested loops leading to incremental 4D-VAR; e.g. *Courtier et al*. [1994]) to cut down the numerical cost without sacrificing too much on the basic analysis method itself. Individually examining these algorithms and the resultant flavors of 4D-VAR is beyond the scope of this chapter and a more detailed review of these algorithms may be found in *Talagrand* [2010].

Chapter 4

Background Error Covariance for Atmospheric CO_2 Data Assimilation

4.1 INTRODUCTION

Atmospheric CO_2 observations provide a powerful constraint on the net sources and sinks of CO_2 as well as their spatial and temporal distribution (e.g. *Le Quéré et al*. [2009]). The advent of several remote-based instruments for observing CO_2 is expected to provide better insight into the critical controls over the atmospheric CO_2 growth (e.g. *Scholes et al*. [2009]) and potentially improve atmospheric flux inversions (e.g. *Chevallier et al*. [2009b]; *Baker et al*. [2010]). Assessing the information content of remote-sensing datasets, however, is difficult due to sampling limitations and large data gaps in the observations. These data gaps can be caused by geophysical limitations, such as clouds and aerosols (e.g. *Bösch et al*. [2006]; *Tiwari et al*. [2006]), as well as by retrieval uncertainties. Further analysis in the form of statistical mapping (e.g. *Alkhaled et al*. [2008]; *Hammerling et al*. [2012a; 2012b]) or binning and averaging (e.g. *Tiwari et al*. [2006]; *Crevoisier et al*. [2009]; *Kulawik et al*. [2010]) is often necessary in order to create full-coverage maps and maximize the usefulness of the data.

An alternate approach is to use advanced data assimilation (DA) techniques to extract information about global CO_2 distributions from the available observations. Atmospheric CO_2 data assimilation (e.g. *Engelen et al*. [2004]; *Engelen and McNally* [2005]; *Engelen et al*. [2009];

Li et al. [2012]) yields 4-dimensional fields of atmospheric CO_2 concentrations that are statistically consistent with (1) the information provided by the CO_2 observations, and (2) additional sources of information, for example from model estimates of CO_2 fluxes, atmospheric transport etc. These additional sources of information, commonly termed as the *a priori* model state (or the background state) are typically obtained from a short-range forecast valid at the time of assimilation. The background state incorporates knowledge about the mechanistic processes governing the carbon cycle and/or imposes physical or dynamical constraints to the assimilation. An attractive feature of the DA framework (*Engelen et al*. [2009]) is that along with CO_2 observations it also makes it possible to assimilate relevant meteorological variables such as temperature and humidity, which are known to affect the observed radiances from which CO_2 information is derived. The final analysis is a consistent estimate of the atmospheric CO_2 concentrations, which if produced with sufficiently high accuracy, could be subsequently used for carbon flux inversions (e.g. *Chevallier et al*. [2009a]).

In all DA applications, ranging from atmospheric CO_2 or other constituents to the Numerical Weather Prediction (NWP) problem, the observations and the background are weighted (e.g. *Nichols* [2010]) based on the accuracy of the information sources. Like any other information source, the background state is prone to errors, which are accounted for through the so-called background error covariance statistics (e.g. *Bannister* [2008a]). It is now well accepted that a realistic representation of the background error distribution is highly important for successful data assimilation. The background error statistics filter and spatially spread the information provided by the observations and impose correlation between different model variables. The critical role played by the background has been demonstrated by *Cardinali et al*. [2004], where it was shown that only 15% of the information content of a well-balanced

meteorological analysis is attributable to the observations assimilated during an analysis; the remaining 85% of the information is provided by the background.

A conceptual definition of background error is relatively trivial, in the sense that it corresponds to the difference between the background atmospheric state and the true atmospheric state. Nonetheless, realistic estimation of the error statistics is not straightforward for several reasons. First, the true atmospheric state is never exactly known, such that the background errors and associated covariances must be estimated from surrogate data. A second difficulty is that error contributions to the background are relatively complex, because the background is the result of a complex data assimilation procedure that involves interplay between the observations assimilated in the past, the analysis formulation, and a forecast model operator. A third difficulty is that the size of the background error covariance matrix is too large to be stored explicitly, necessitating a variety of reduced-rank approximations to make the computations feasible.

Various techniques have been developed to tackle these challenges (*Bannister* [2008a; 2008b]), and there is now a substantial literature on background error statistics for NWP and constituent assimilation applications, covering their nature, estimation, and practical implementation in operational settings (e.g. *Derber and Bouttier* [1999]; *Chapnik et al*. [2004]; *Fisher* [2004]; *Deckmyn and Berre* [2005]; *Zagar et al*. [2005]; *Belo Pereira and Berre* [2006]; *Berre et al*. [2006]; *Benedetti et al*. [2007]; *Buehner and Charron* [2007]; *Constantinescu et al*. [2007]; *Pannekoucke et al*. [2007; 2008]; *Raynaud et al*. [2009]; *Berre and Desroziers* [2010]; *Boloni and Horvath* [2010]; *Bonavita et al*. [2010]; *Cheng et al*. [2010]; *Hess* [2010]; *Michel et al*. [2010]; *Giuseppe et al*. [2011]; *Singh et al*. [2011]; *Varella et al*. [2011]; *Brousseau et al*. [2012]; *Massart et al*. [2012]).

The proposed approaches are problematic for atmospheric CO_2 data assimilation, however, because they fail to take into account: (a) the significant uncertainty associated with the surface CO_2 fluxes, and (b) the errors associated with the CO_2 transport model. Forecast or ensemble-based techniques prescribe the background statistics based only on the internal uncertainties in the transport model, but these are not a realistic representation of the true transport model errors. Additionally spatial and temporal variations in the CO_2 fluxes are a key driving force behind atmospheric CO_2 distributions, and consequently the measured CO_2 concentrations. Failing to capture the true magnitude of the transport errors and ignoring the uncertainty of the fluxes results in an under-estimation of the background error statistics, which makes the assimilation reject the observations in favor of the background. This problem is especially severe during the Northern Hemisphere summer when the uncertainties in the background CO_2 fluxes are expectedly high. In such cases, one has to artificially inflate the background error (*Engelen et al*. [2009]), but this strategy still does not account for the spatial error correlations due to the flux uncertainties.

The primary goal of this study is to outline an approach for parameterizing the background error covariance for atmospheric CO_2 data assimilation, in a way that includes the statistics of errors resulting from both flux and transport uncertainties. The proposed approach is based on the assumption that the difference between modeled CO_2 concentrations (henceforth, denoted as ΔCO_2) from two state-of-the-art global models is statistically representative of the background errors. Because any two models provide a limited sample of the true background error distribution, it is beneficial to ensure that these models are different both in terms of the underlying fluxes and the transport fields driving them to capture realistic variability in the

background error. The error statistics can then be generated from the ΔCO_2 fields using spatial statistical tools.

In this study, the resulting background error statistics are implemented within the atmospheric CO_2 4D-VAR system at ECMWF. This system, as described in *Engelen et al*. [2009], is akin to a NWP-DA setup, in which CO_2 mixing ratios are constrained along with other atmospheric variables such as temperature, winds, surface pressure, and humidity, to obtain a consistent estimate of the atmospheric CO_2 concentrations. Experiments are designed to evaluate: (a) whether the difference between two models can be used as a proxy for the statistics of the background errors, (b) the extent to which the representation of background errors has a discernible impact on CO_2 estimates, and (c) whether including realistic statistics of errors improves the performance of the data assimilation system, relative to simply accounting for internal transport model uncertainties as available from a forecast-based technique (standard version of the NMC-method, according to the terminology of *Široká et al*. [2003]). The predictions of the 4D-VAR analyses are evaluated using independent observations of CO_2 from aircraft profiles and column-averaged dry mole fractions of CO_2 (i.e., XCO_2) from the Total Column Carbon Observing Network (TCCON).

Although the application presented here primarily focuses on atmospheric CO_2 data assimilation, the proposed approach is relevant for other trace gas assimilation applications, especially ones in which the background errors are influenced by both atmospheric transport and emission patterns. Additionally, this study provides insights into improvements that can be achieved within operational atmospheric data assimilation systems from better representations of background error statistics in general.

4.2 EXPERIMENTAL FRAMEWORK

4.2.1 Four-dimensional variational data assimilation

The atmospheric 4D-VAR data assimilation system used in this study is based on the one described by *Engelen et al* [2009]. It uses the ECMWF Integrated Forecasting System (IFS) transport model with CO_2 fluxes prescribed at the surface based on climatological and inventory data. The system assimilates the same meteorological observations as the operational ECMWF system, in addition to observations constraining CO_2. We avoid repeating the details of the data assimilation system here, and instead point out the two major changes relative to *Engelen et al*. [2009]. The specification of the background covariance matrix, however, forms the primary focus of this study and is discussed separately in Section 4.2.2.

First, the IFS version has been updated since the study of *Engelen et al*. [2009]. The version used in this study is based on CY37r3 (ECMWF, IFS documentation CY37r2, 2011, http://www.ecmwf.int/research/ifsdocs/CY37r2/index.html), which became operational in November 2011, and incorporates several improvements to the atmospheric forecast model and the data assimilation system at ECMWF.

Second, the focus in *Engelen et al*. [2009] was on the assimilation of radiances from the Atmospheric Infrared Sounder (AIRS; *Chahine et al*. [2006]) and the Infrared Atmospheric Sounding Interferometer (IASI; *Crevoisier et al*. [2009]) instruments. Instead, this study uses L2 retrieval data from the Greenhouse gases Observing SATellite "IBUKI" (GOSAT; *Kuze et al*. [2009]; *Yokota et al*. [2009]). These datasets are obtained from version 2.9 of the ACOS algorithm (e.g. *O'Dell* et al. [2012]; *Crisp et al*. [2012]). Based on the recommendations of the ACOS team, only the high (H) gain observations with the master quality flag equal to "good" are used in the assimilation process. Although first estimates for bias correction of the ACOS data do

exist (*Wunch et al*. [2011a]), these are not used in the experiments in this study. GOSAT data provide a stronger constraint on CO_2 near the surface relative to AIRS and IASI observations, but have relatively poor geographical coverage.

4.2.2 Specification of the background error covariance matrix

4.2.2.1 NMC method

The background error covariance matrix used in *Engelen et al*. [2009] is based on the NMC (National Meteorological Center, nowadays named National Center for Environmental Prediction) method (*Parrish and Derber* [1992]). The NMC method is also based on the principle of using a surrogate quantity to represent the background errors, where the surrogate is typically chosen to be the differences between forecasts of different length valid at the same time. In its simplest form (e.g. *Široká et al*. [2003]), the NMC method is implemented by taking the differences between 24 h forecasts and 12 h forecasts over a one-month period. The main advantage of this method is that the forecasts required for calculating the background error statistics are readily available during the DA process.

One limitation of this approach is that the variance of background errors in data sparse regions is underestimated (e.g. *Berre et al*. [2006]) because any differences between forecasts of different lengths are attributable to information from observations within the period between the starting times of the two forecasts. A second limitation is that this approach does not account for the uncertainty associated with the underlying fluxes because the two forecasts start from the same set of prescribed fluxes. This also implies that if there is no significant seasonality in the observational constraint then the error statistics defined via the NMC method remain invariant in time, and will fail to account for the seasonal variability in the background errors. Third, while the NMC method is able to account for internal (i.e. within the same model) transport

uncertainties, they cannot represent the epistemic errors due to the structure of the transport model itself. Finally, the implicit formalism of the NMC method (i.e., the analysis of increments) emphasizes the small-scale structures in the background CO_2, missing much of the large-scale variability. Interestingly, this is different from NWP-related applications where it has been pointed out (e.g. *Berre et al*. [2006]; *Belo Pereira and Berre* [2006]; *Storto and Randriamampianina* [2010]) that the NMC method likely overestimates the error correlations. This difference is purely based on the different scales at which the chaotic weather dynamics and the atmospheric CO_2 processes operate.

4.2.2.2 Δ-statistics method

Analogous to the NMC method, the Δ-statistics are formulated using a surrogate quantity, which is chosen to be the difference in the modeled CO_2 concentrations (ΔCO_2) from two global models. In this study, the two selected models are PCTM-GEOS4 CASA (*Kawa et al*. [2004]) and IFS-ORCHIDEE (e.g. *Hollingsworth et al*. [2008]; *de Rosnay* et al. [2003]; *Krinner et al*. [2005]), which are based on different biospheric CO_2 flux distributions (CASA vs. ORCHIDEE) propagated by different atmospheric transport (PCTM-GEOS4 vs. IFS). These models also have distinct specifications of ocean fluxes as well as fossil fuel and fire emissions.

The ΔCO_2 fields (Figure 4.1) are obtained 3-hourly at a horizontal resolution of 1° by 1.25° for 60 sigma-hybrid levels of the atmosphere. The horizontal and vertical error correlations for generating the background error covariance matrix is based on these ΔCO_2 fields. The vertical error correlations distribute the information of the observations in the vertical because they represent the correlation between errors at different altitudes. In our case, the correlation coefficients of ΔCO_2 between different atmospheric (or model) levels are constant in space but vary monthly to capture the seasonality in the ΔCO_2 field.

The horizontal error correlations define how errors are correlated between grid boxes, and define the degree to which CO_2 is adjusted around grid boxes containing observations. In our case, the horizontal error correlations are obtained from a variogram analysis, a quantitative tool in geostatistics that has been successfully used to characterize the spatial and temporal structure of atmospheric CO_2 (e.g. *Alkhaled et al*. [2008]; *Hammerling et al*. [2012a]). The horizontal correlations are themselves spatially variable, and are defined separately for each atmospheric (or model) level. Separate horizontal error statistics are calculated for each atmospheric (or model) level due to the difference in the patterns of CO_2 gradients at different levels of the atmosphere (Figure 4.1). Near the surface (Figure 4.1A and B), differences between the

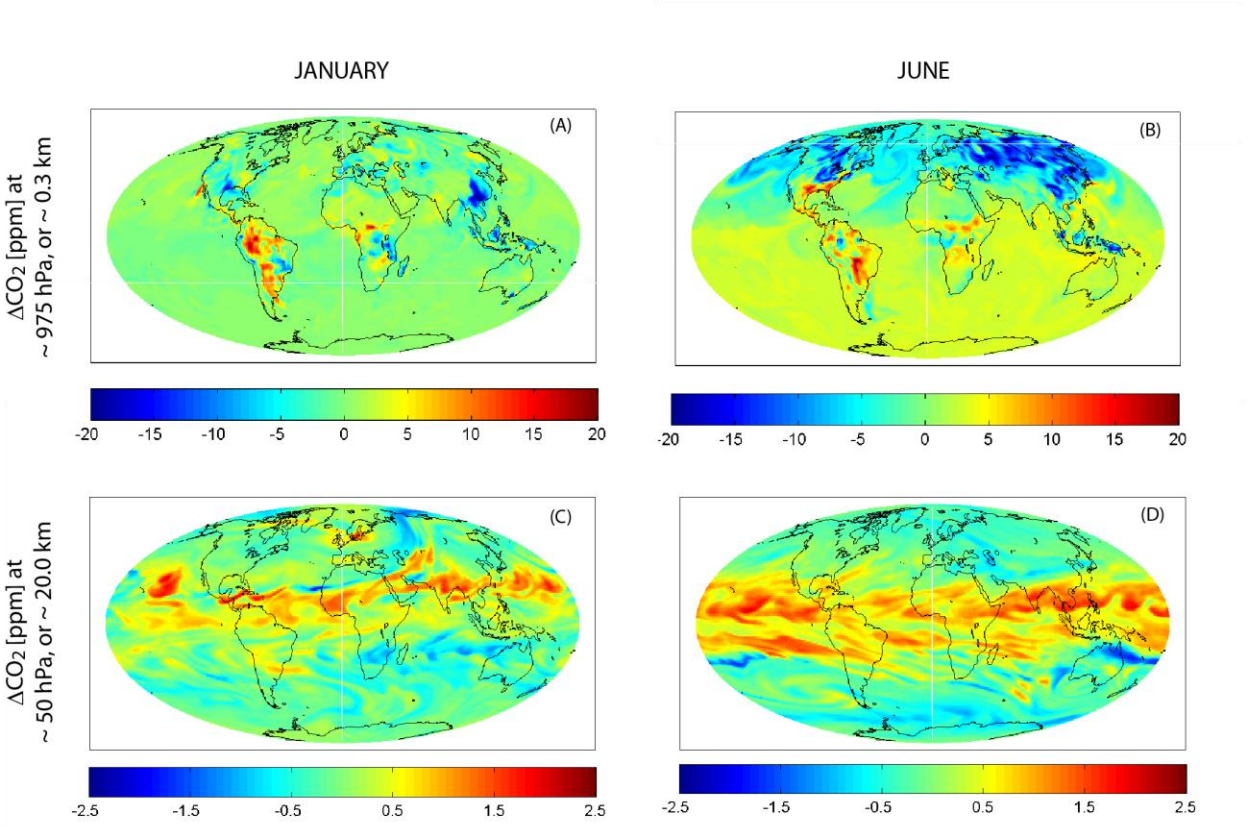


Figure 4.1- *ΔCO_2 fields for: (a) January (A and C) and June (B and D), and (b) at two different atmospheric levels, 975 hPa or ~ 300 m (Panels A and B) and 250 hPa or ~ 10 km. The fields are shown for the 15th of the month at 1800 h UTC.*

prescribed fluxes, result in higher errors in the background. In addition, the interaction between boundary layer dynamics and biospheric emissions (a.k.a. 'the rectifier effect, *Denning et al*. [1999]) also contributes to large variability in the CO_2 concentrations near the surface. Conversely, in the free troposphere (Figure 4.1C and D), CO_2 is more dispersed and well-mixed, yielding smaller and smoother errors. Note that in Figure 4.1 the ΔCO_2 fields are shown for two typical months - January and June 2010, which are representative of Northern Hemispheric winter and summer, respectively. As can be seen in this figure, the seasonality in the ΔCO_2 fields is more evident near the surface (Figure 4.1A and B) relative to higher levels in the atmosphere (Figure 4.1C and D).

The plots in Figure 4.1 also show that within each atmospheric (or model) level, significant regional variability exists in the ΔCO_2 fields. These reflect the regional differences in surface fluxes between the models, as well as differences due to different representations of global atmospheric transport. Previous work by *Alkhaled et al*. [2008] and *Hammerling et al*. [2012a] using column-averaged CO_2 concentrations (i.e., XCO_2) highlighted that significant regional variability exists in the spatial covariance structure of atmospheric CO_2 concentrations, and hence any spatial analysis needs to be done regionally rather than globally. Keeping these features in mind, the spatial variability ($\gamma_{region}(d)$) at a particular model level is defined as:

$$\gamma_{region}(d) = \frac{1}{2}\left[(\mathbf{\Delta CO_2}(x_{region}) - \mathbf{\Delta CO_2}(x_{region+d}))^2 \right] \quad 4.1$$

where x_{region} and $x_{region+d}$ denote the differences in modeled CO_2 over two locations separated by a distance d. Similarly to *Alkhaled et al*. [2008], here: a) regions are defined as overlapping 2000km radius circles, centered at the model grid cell, and b) $\gamma_{region}(d)$ is constructed using pairs of points, with the first point being within the specified region (ΔCO_2 (x_{region})) and the other

point being either within or outside that region (ΔCO_2 ($x_{region+d}$)). The horizontal variability, thus, accounts for both the observed variability within each sub-region (by using all available pair of points within a region) and large scale variability (by using a random sample of the points outside the region). Once the regional variability of ΔCO_2 is defined, it can be represented using an exponential decay in the spatial correlation as a function of separation distance (d), parameterized by a variance (σ^2) and a correlation length ($3l$), and a positive-definite function describing the decay of auto-correlation as the separation distance (d) increases:

$$\mathbf{Q}(d) = \sigma^2 - \gamma_{theo}(d) = \sigma^2 \exp\left(-\frac{d}{l}\right) \qquad 4.2$$

Fitting the correlation length and the variance parameters is possible using a simple or weighted least-squares technique, or more computationally efficient minimization schemes such as the limited memory BFGS (L-BFGS; *Nocedal and Wright* [2006]) as used in this study. The horizontal error correlation coefficients are now obtained as the ratio of $\mathbf{Q}(d)/\ \sigma^2$.

Theoretically, both the horizontal and the vertical error correlations can be generated for each time period (i.e. 3-hourly) when the modeled CO_2 concentrations are available. Based on our understanding of the surface CO_2 exchange, the atmospheric structure and the transport of air masses, significant seasonal (or monthly) differences in the background error are expected. At daily or sub-daily time scales, however, the differences in the background error are local and mostly remain confined within the boundary layer. Hence in this study, these error correlations are generated once each month (on the 15th day at 1800 UTC) that is assumed to be representative of the typical variability that would be observed during an individual day in the month.

The properties of the background error covariance inferred from the Δ-statistics are subsequently modeled using the mathematical framework of wavelet-like, non-orthogonal basis

functions having simultaneous localization both in space and wave number. This mathematical formulation of the background error covariance matrix is based on *Fisher* [2004; 2006] and the application to tracer variables is described in *Benedetti and Fisher* [2007].

4.3 SAMPLE APPLICATION

4.3.1 Experiments

GOSAT observations for the year 2010 are assimilated into the 4D-VAR system to generate global 4D distributions of atmospheric CO_2. Two independent sets of experiments are run, in which the background error covariance is prescribed based on the Δ-statistics (henceforth "*analysis with the Δ-statistics*") and the NMC approach (henceforth "*analysis with the NMC statistics")*, respectively. While the background error using the Δ-statistics method is defined as outlined in Section 4.2.2.2, the NMC method typically under-estimates the background error. In order to offset this under-estimation, following *Engelen et al*. [2009] the standard deviations of the background errors from the NMC-based statistics are inflated by a factor of 8 with a maximum of 15 ppm. The inflation factor is based on comparisons between the CO_2 model concentrations and independent surface and aircraft observations. All other parameters in these experiments are held the same to allow a straightforward evaluation of the impact on the CO_2 estimates due to different parameterization of the background errors statistics.

A third set of experiments is run in which fluxes are prescribed but no observational constraint on CO_2 is provided (henceforth "*unconstrained model run")*, to assess the impact of the GOSAT observations on the assimilation and the impact of each of the background error representations. This experiment uses only meteorological observations and transports the CO_2 starting from the same initial field on 1 January 2010 as the other two sets of experiments.

4.3.2 Evaluation of the 4D CO_2 fields

Given that the true atmospheric CO_2 concentrations are unknown globally, the 4D-VAR estimates from the three experiments are evaluated using two sets of independent observations (Figure 4.2) that are not included in the assimilation process – column-averaged dry mole fractions of CO_2 (i.e., XCO_2) from the Total Column Carbon Observing Network (TCCON) and vertical profiles of CO_2 from aircraft observations. Both these observational datasets have much higher accuracy than the 4D-VAR analysis, and are assumed to be representative of the true atmospheric state.

TCCON is a global network of calibrated ground-based Fourier transform spectrometers that measure the total column amount of various species (including CO_2) by recording the direct solar spectra in the near-infrared spectral region (*Wunch et al*. [2011b]). The TCCON observations are compared to column-averaged 4D-VAR analysis estimates within six latitudinal bands (90-60 N, 60-30 N, 30-0 N, 0-30 S, 30-60 S and 60-90 S) to assess the impact of latitudinal differences in the background error statistics on the 4D-VAR analysis.

Mean absolute errors (MAE; e.g. *Willmott and Matsuura* [2005]) are calculated across all TCCON sites within each latitudinal band, as is the uncertainty of the mean:

$$\mathrm{MAE} = \frac{1}{\sum_{i=1}^{\mathrm{M}} \mathrm{N}_i} \left[\sum_{i=1}^{\mathrm{M}} \sum_{j=1}^{\mathrm{N}_i} \left| \hat{\mathrm{y}}_{m_i,n_j} - \mathrm{y}_{m_i,n_j} \right| \right] \tag{4.3}$$

$$\text{uncertainty of MAE} = \sqrt{\frac{1}{\left(\sum_{i=1}^{\mathrm{M}} \mathrm{N}_i \right)^2} \left[\sum_{i=1}^{\mathrm{M}} \sum_{j=1}^{\mathrm{N}_i} \left| \hat{\mathrm{y}}_{m_i,n_j} - \mathrm{y}_{m_i,n_j} \right|^2 \right]} \tag{4.4}$$

where M represents the total number of TCCON sites, N_i represents the total number of observations at the i^{th} TCCON site, y_{m_i,n_j} is the j^{th} TCCON observation at the i^{th} site, and $\hat{y}_{m_i,n_j}$ is

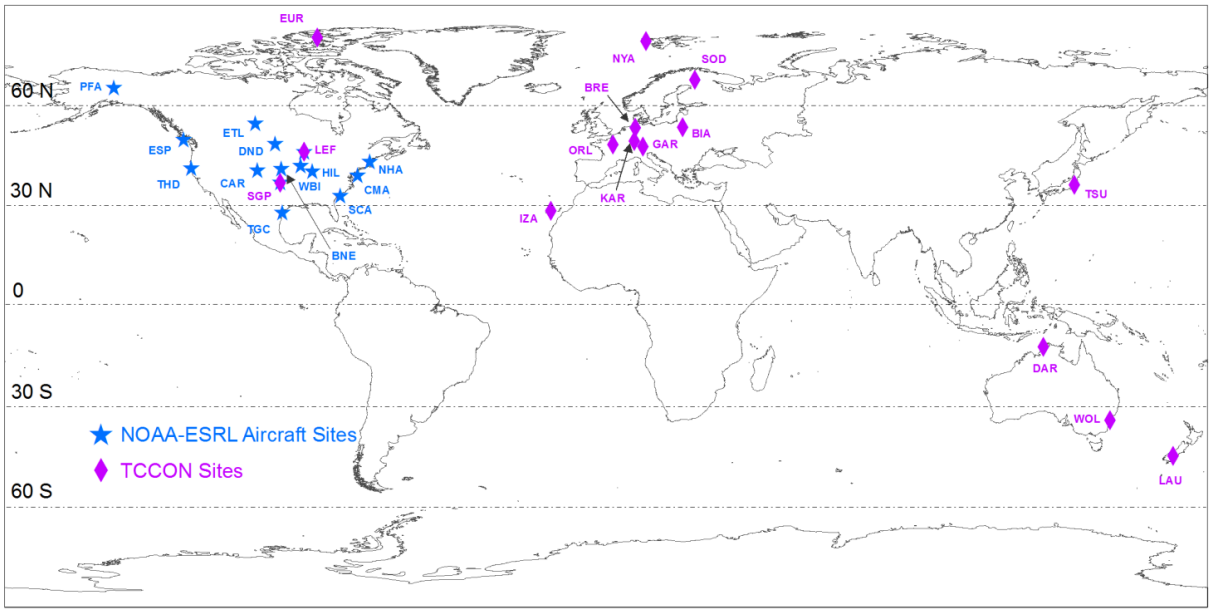


Figure 4.2*- Location of the aircraft and TCCON sites. The three-letter codes for the sites are defined in Appendix B.*

the corresponding 4D-VAR analysis for that TCCON observation. Additionally, the MAE for each TCCON site is also calculated as:

$$\mathrm{MAE}_i = \frac{1}{\mathrm{N}_i}\left[\sum_{j=1}^{\mathrm{N}_i}\left|\hat{y}_{m_i,n_j} - y_{m_i,n_j}\right|\right] \qquad 4.5$$

The second evaluation is carried out using aircraft observations available over North America. The National Oceanic and Atmospheric Administration Earth System Research Laboratory (NOAA-ESRL) has been conducting long term aircraft monitoring (http://www.esrl.noaa.gov/gmd/ccgg/aircraft/), in which vertical profiles of various trace gases (including CO_2) are observed in the troposphere (i.e., surface to 8 km altitude) with high accuracy (e.g. *Tans* [1996]; *Crevoisier et al*. [2010]). Sampling frequencies are weekly or biweekly for most sites. The aircraft observations and the 4D-VAR analysis estimates are divided into four altitude bins, 0-2 km, 2-4 km, 4-6 km and 6-8 km. For each altitude bin, Equations 4.3-4.5 are applied to calculate the MAE metrics between the 4D-VAR analysis

estimates and the aircraft observations. This evaluation exercise, thus, provides quantitative information regarding the impact of the background error statistics in the vertical. Because aircraft observations are available only over North America for the study period, only a single latitude band between 90 – 0 N is considered for this second evaluation exercise.

4.4 RESULTS

4.4.1 Background error correlation from the Δ-statistics

Figure 4.3 shows the horizontal error correlation length and the variance for the near-surface ΔCO_2 fields shown in Figure 4.1A and B, showing both regional and seasonal differences in the inferred covariance parameters. In June, strong biospheric signals dominate the Northern Hemisphere and the overall flux uncertainty and variability is higher relative to January. Consequently the horizontal error correlations in Figure 4.3 show higher variance and shorter correlation length over the Northern Hemisphere. Analyses at higher atmospheric (or model) levels indicate longer correlation lengths and lower variance. This is expected because the impact of the near-surface CO_2 exchange and transport is dampened higher in the atmosphere, and the variability is dominated by the synoptic-scale mixing of air masses.

The vertical error correlations based on the ΔCO_2 fields are stronger relative to those derived from the NMC analysis, show distinct seasonality, and decay more slowly across levels (Figure 4.4). As a result, analysis based on the Δ-statistics will spread the information from the observations more in the vertical relative to an analysis using NMC-statistics. In the upper troposphere and the stratosphere (Figure 4.4B) the vertical error correlations from the two methods become more similar to one another, although the vertical error correlations from the Δ-

statistics still decays more gradually compared to the NMC-statistics (Figure 4.4B – between 20 - 0 km)

If the ΔCO_2 fields are a reasonable representation of the structure of the background errors, then its corresponding horizontal and vertical error correlations indicate that the background error statistics should vary spatially and temporally. By specifying these error statistics to be constant over large areas or forcing them to be invariant in time, the NMC statistics may underestimate (or over-estimate) the magnitude of the true errors and therefore degrade the DA analysis. These error statistics also provide an initial indication of the observational impact on the analysis. The long correlations are a result of the large scale structures in the CO_2 background. Within a DA analysis, specifying these long error correlation functions will make the analysis filter out any small-scale features while propagating the information from the observations to large distances, in the horizontal and the vertical.

4.4.2 Impact on the 4D-VAR analysis due to the background error statistics

A sample of atmospheric CO_2 fields near the surface from the three experiments is shown in Figure 4.5. All three experiments capture the latitudinal and inter-hemispheric gradients in CO_2 as well as the seasonal cycle of the Northern Hemisphere CO_2, i.e., buildup of CO_2 concentrations in January when the majority of the terrestrial biosphere is dormant and a significant drawdown of CO_2 in June due to the vegetation uptake.

A closer evaluation of the two 4D-VAR analysis using the GOSAT observations indicates that the difference between the two analysis estimates are greater near the surface, where most of the variability in the CO_2 processes (such as fossil fuel emissions, biospheric exchanges, influence of boundary layer variations etc.) are present.

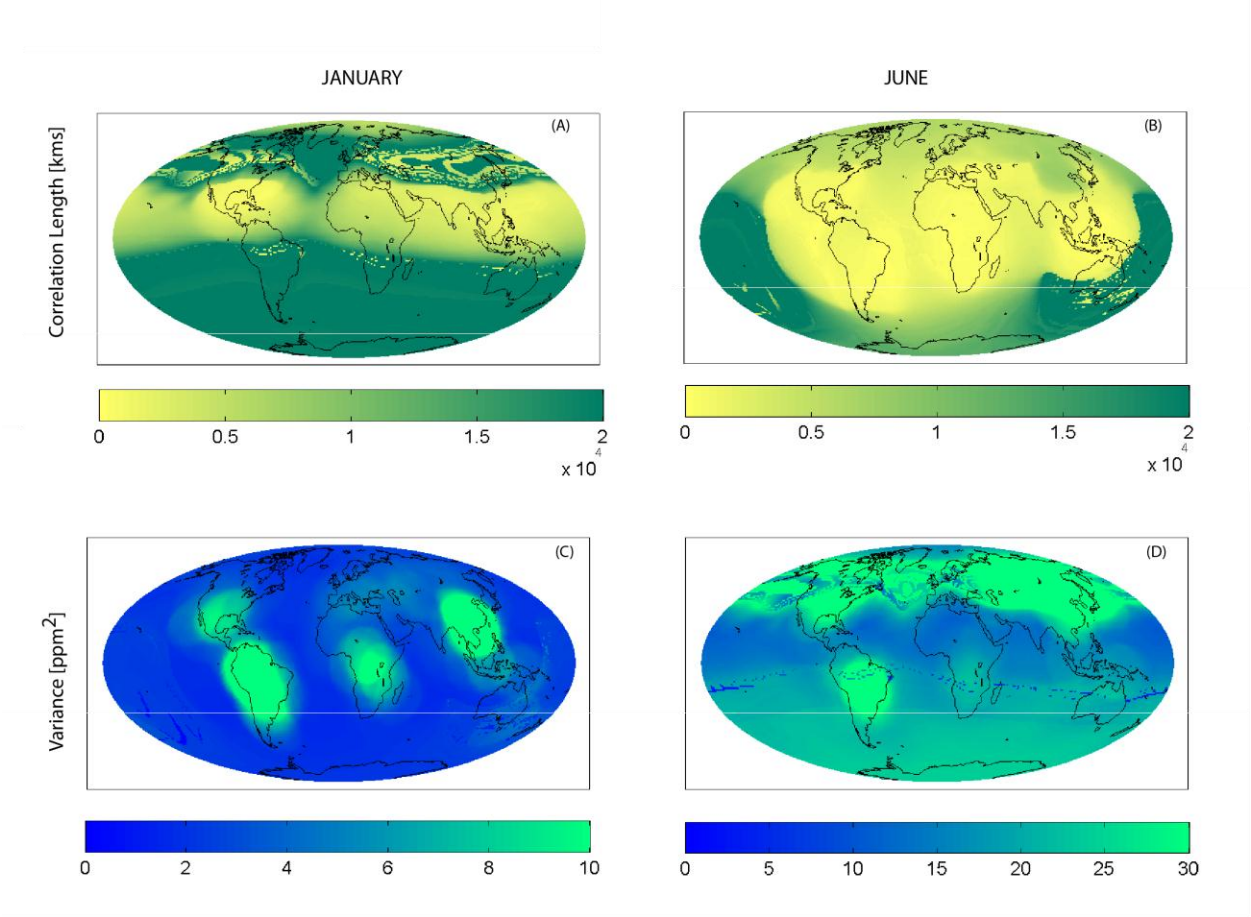


Figure 4.3- *Inferred correlation length (3l) and variance (σ^2) parameters for January (A and C) and June (B and D). The covariance parameters are shown for the 15th day of the month, 1800h UTC at 975 hPa (i.e., ~ 0.3 km). Note the higher variance for June is indicative of greater variability in the surface CO_2 processes during this month.*

Relative to the background, the analysis with the Δ-statistics allows more adjustment by the GOSAT observations than the analysis using the NMC-statistics. A typical example is visible over Eurasia (compare the CO_2 fields in Figure 4.5B and 4.5D with those in 4.5F). The NMC-statistics tends to produce localized changes only while the adjustments made by the Δ-statistics are over larger scales. The mean absolute difference in CO_2 concentrations between the analysis with the NMC statistics and the unconstrained model run (January – 0.22 ppm, June – 0.36 ppm) is smaller and less variable across seasons than the corresponding difference between the analysis with the Δ-statistics and the unconstrained model run (January – 0.63 ppm;

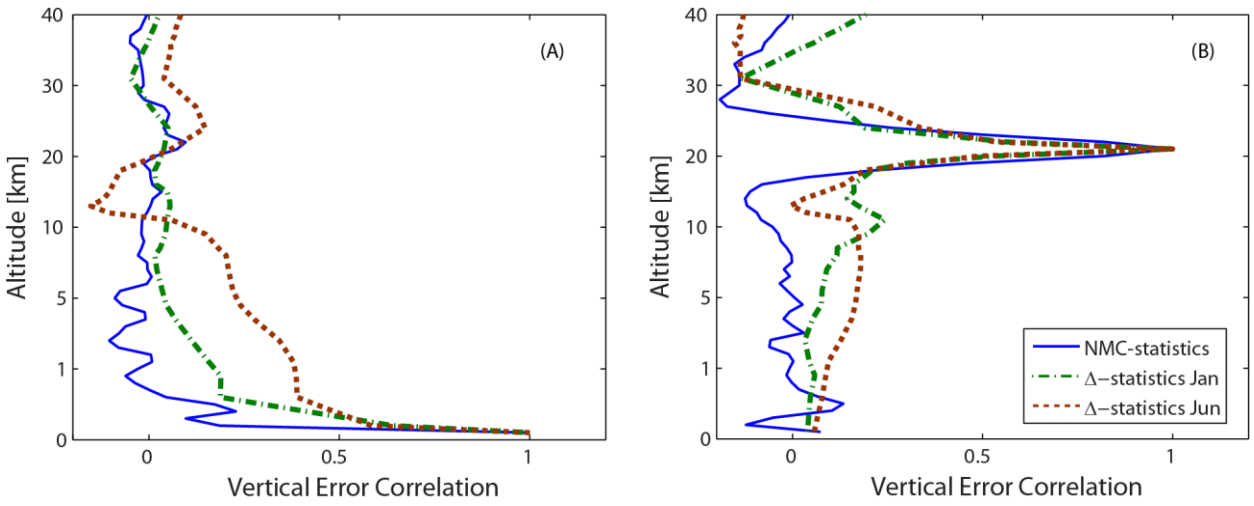


Figure 4.4*- Vertical error correlations from the seasonally-static NMC-method and the Δ-statistics at the surface (A), and at ~50 hPa or ~ 20.0 km (B).*

June - 3.3 ppm). These differences are limited to the lower levels of the atmosphere where: a) the error statistics from the NMC and the Δ–method are most different, and b) the GOSAT observations tend to have the largest impact on the analysis. This supports our original hypothesis that the longer error correlation length (both in the horizontal and the vertical) in the Δ-statistics allows the GOSAT observations to have a greater impact. The relatively 'stiff' or the smaller error variances associated with the NMC NMC-statistics gives less weight to the GOSAT observations in favor associated with the keeping the analysis close to the background CO_2 distribution.

An alternate way of examining the impact of the GOSAT observations is by looking at the CO_2 analysis increments (analysis minus background values), which show the direct impact of the observations on the CO_2 field. As described in *Fisher* [2002], within the ECMWF 4D-VAR system these increments are calculated using the differences between the observations and the model background (short forecast), the observation errors and the background errors. Figure 4.6 shows the increments in total column mean mixing ratio of CO_2 for June 2010 for the NMC experiment and the Δ-statistics experiment. The CO_2 increments in the NMC experiment

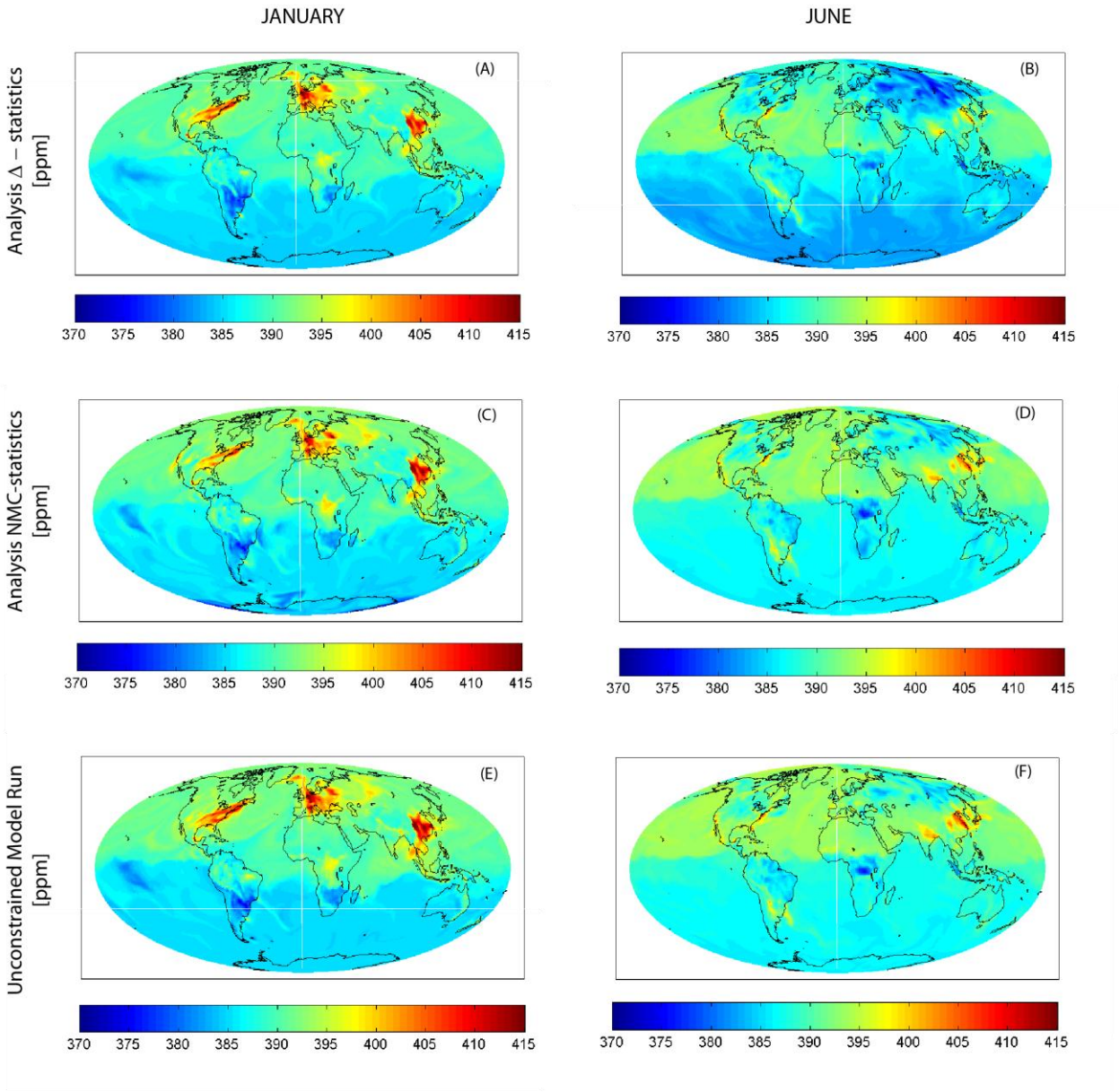


Figure 4.5- *CO_2 concentrations from the 4D-VAR analysis based on the two background error statistics – Δ-statistics (A and B), NMC statistics (C and D) and the no-assimilation model run (E and F) for January and June 2010. The plots are shown for the 15th day of the month, 1800h UTC at 975 hPa (i.e., ~ 0.3 km).*

are generally small-scale, while the Δ-statistics allow much broader scale adjustments to the CO_2 fields. This follows from the longer error correlation lengths and higher variance prescribed in Δ-statistics compared to the NMC-statistics. The accuracy of these adjustments is next evaluated using independent observations.

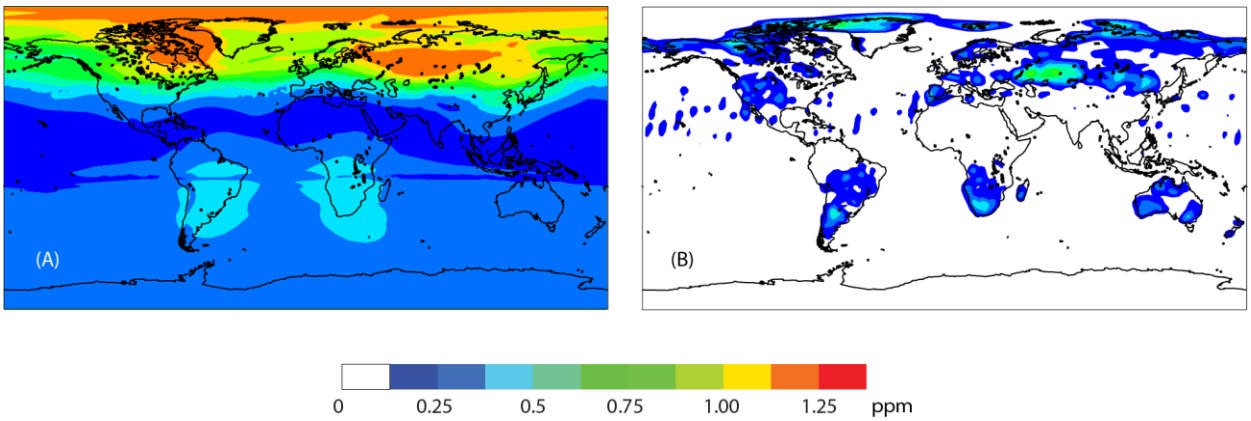


Figure 4.6*- Column-averaged CO_2 mixing ratio analysis increments based on the two background error statistics –Δ-statistics (A), and NMC-based statistics (B), averaged over the month of June 2010.*

4.4.3 Evaluation of the 4D-VAR analysis

The quality of the 4D-VAR analysis is assessed by comparing the atmospheric CO_2 fields to observations from the aircraft and the TCCON networks. Figures 4.7A and 4.7B show example profiles collected on 9th Jan 2010, and 26th June 2010, over Worcester, Massachusetts (site code-NHA in Figure 4.2). In both cases, the 4D-VAR analysis reduces the mismatch between the background and the true atmospheric CO_2 state. The ability of the GOSAT observations to adjust the details of the background CO_2 profile, however, is limited by the stiff NMC-method relative to the more flexible Δ-statistics approach. Figures 4.8A and 4.8B show the TCCON observations collected over Bialystok, Poland (site code – BIA in Figure 4.2) for different days in January and June 2010. The NMC statistics again limit the degree to which the GOSAT observations can adjust the background state.

Overall, the CO_2 estimates based on the Δ-statistics show significant improvement relative to the analysis using NMC statistics during the Northern Hemisphere summer, when significant uncertainty in the fluxes are present (Figure 4.9). The variability of the MAE across TCCON sites is also smaller when using the Δ-statistics approach. This denotes that the CO_2

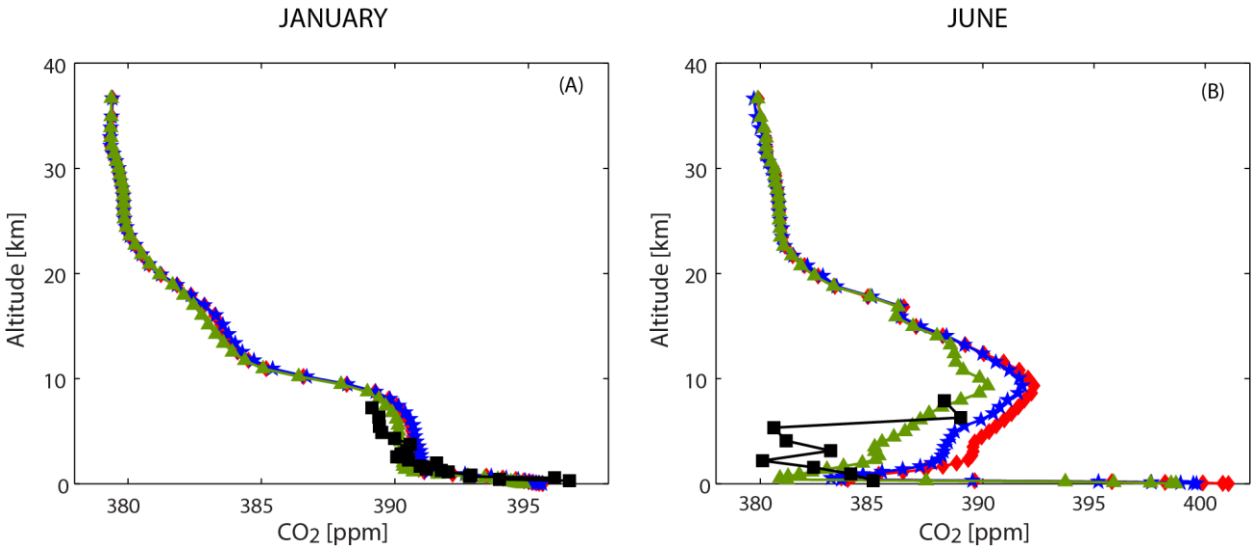


Figure 4.7*- Evaluation of 4D fields of CO_2 using aircraft observations over Worcester, Massachusetts (site code – NHA) on 9th January 2010 (A), and 26th June 2010 (B).*

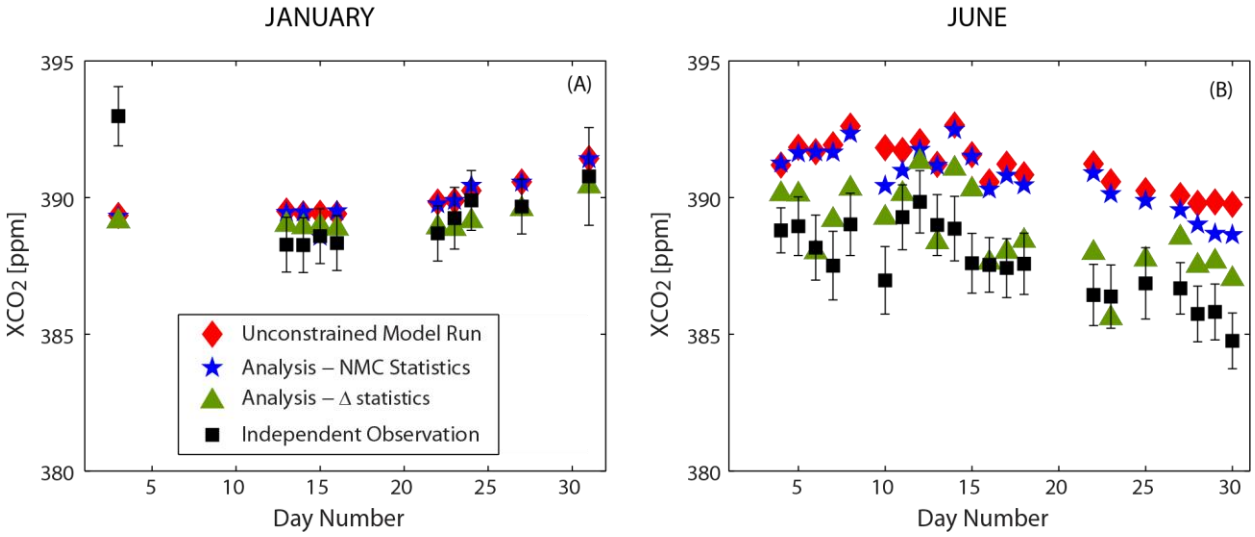


Figure 4.8*- Evaluation of column-averaged CO_2 using TCCON observations over Bialystok, Poland (site code – BIA) during January 2010 (A), and June 2010 (B).*

concentrations obtained from the analysis with the Δ-statistics are consistently closer to the observed CO_2 from the TCCON sites. During the Northern Hemisphere winter, the analysis with the Δ-statistics performs similarly to the analysis using NMC-based estimates (Figure 4.9A, 60-30 N) or slightly worse (Figure 4.9B, 30-60 S). The performance over the Southern Hemisphere, however, is difficult to judge given the dearth of TCCON sites.

The aircraft profiles allow us to also examine the assimilation performance at different levels of the atmosphere. As seen from Figures 4.7A and 4.7B, at higher levels of the atmosphere (~ ≤ 50 hPa or ≥ 20 km), the differences in the analysis estimates become negligible due to: (a) the background error covariance from the Δ-statistics and the NMC-statistics being similar, and (b) the GOSAT observations being uninformative. Both analysis estimates tends to remain close to the background model CO_2. Between 10 and 20 kms (i.e., ~ 200 hPa and 50 hPa), however, occasional differences are visible between the 4D-VAR analysis and the background.

This pattern is also generally true across aircraft sites (Figure 4.10). Given that the GOSAT sensitivity to CO_2 is very low at this level, any information from the observations can only change CO_2 concentrations at this level through transport or through the information spreading of the background covariance matrix. The Δ-statistics spread information more in the vertical relative to the NMC-statistics. Whereas the MAE for the analyses with the Δ-statistics and the NMC statistics are similar higher up in the atmosphere (Figure 4.10A and 4.10B, 6-8 km), differences in the MAE are more clearly visible lower down (Figure 4.10A and 4.10B, 0-2 and 2-4 km). Similar to the TCCON evaluation, the analysis with the Δ-statistics outperforms the other analysis using the NMC statistics and the unconstrained run slightly during the summer, but does slightly worse (Figure 4.10A) during the winter. Recall that almost all of the aircraft sites are located over North America, where there is little variability in the fluxes for the winter months due to reduced biospheric activity.

Before concluding this section it is worthwhile to discuss a couple of caveats regarding the use of background statistics. Since the 4D-VAR configuration assumes unbiased Gaussian statistics, any systematic differences between the observations and the model background may be interpreted incorrectly. When the background error is small relative to the observation error, as is

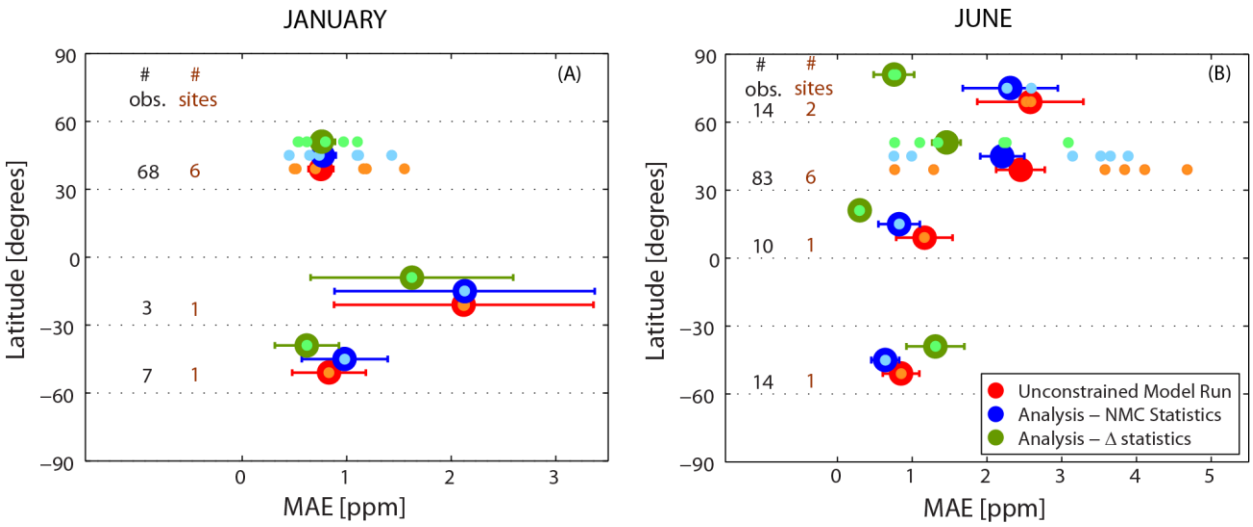


Figure 4.9*- MAE between the column-averaged CO_2 estimates from the 4D-VAR analysis and the TCCON observations, binned by latitude for January 2010 (A), and June 2010 (B). MAE between the 4D-VAR analysis and observations from each of the TCCON sites are shown as the smaller horizontal dots. The error bars on the MAE take into account the total number of independent observations (# obs.) available within each latitude band, which are reported in black text. Values reported in brown denote the number of TCCON sites (# sites) available for evaluation within each latitude band.*

the case with the NMC experiment, the analysis will be directed towards the model background, irrespective of the bias in either the observations or the model background itself. When the background error is large compared to the observation error, as is the case in our Δ-statistics experiment, the analysis is directed towards the observations in the areas where observations are available. Ultimately this leads to significant differences in the final analysis. Previous experiments (results not shown) with a biased set of AIRS data demonstrated that for the Δ-statistics to provide reasonable results, the observational constraint should be of a fair quality. The reduced performance of the Δ-statistics during the winter months could therefore also be a by-product of biases in the CO_2 retrievals. During the winter months, even the analysis with the NMC-statistics does worse than the unconstrained model run (Figures 4.9A – 30-60 N and 4.10A -0-2 km, 2-4 km), which points to an inconsistency between the constraint provided by the GOSAT data and the available independent observations.

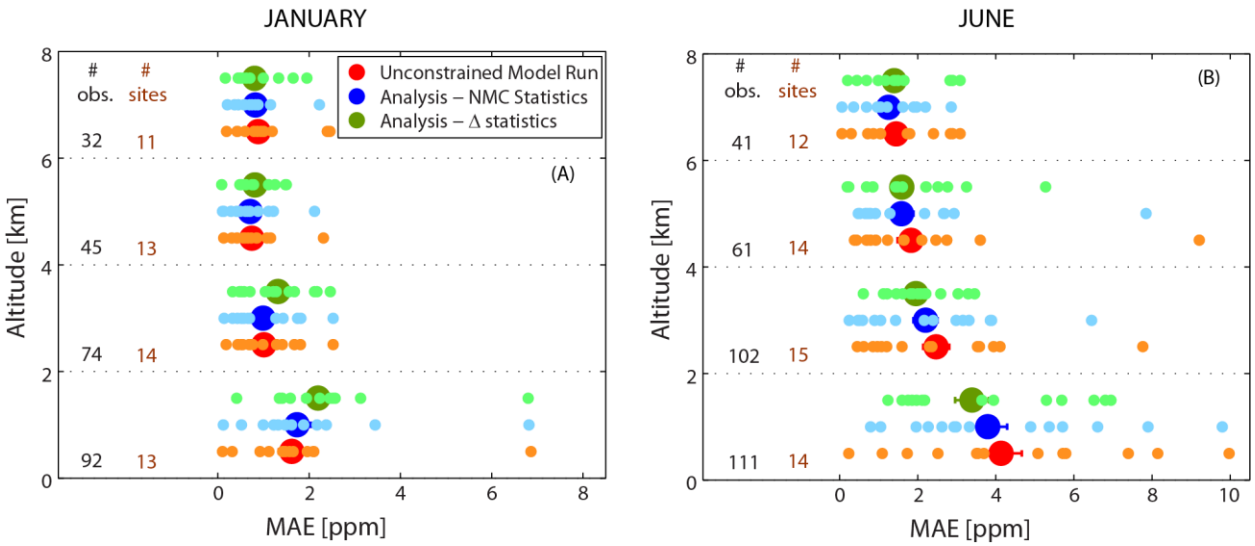


Figure 4.10*- MAE between the CO_2 estimates from the 4D-VAR analysis and the aircraft observations, binned by altitude for January 2010 (A), and June 2010 (B). MAE between the 4D-VAR analysis and observations from each of the aircraft sites are shown as the smaller horizontal dots. The error bars on the MAE take into account the total number of independent observations (# obs.) available within each altitude band, which are reported in black text. Values reported in brown denote the number of aircraft sites (# sites) available for evaluation within each altitude band.*

4.5 SUMMARY

The background error covariance matrix is a critical component of an atmospheric data assimilation system and its realistic estimation is necessary to make efficient use of the observational information. This study examined the specification of the background error statistics for the atmospheric CO_2 data assimilation system in place at ECMWF. Currently, the background error statistics for this system are based on the NMC method, which not only underestimates the background errors but also fails to take into account any spatial and temporal variations in the underlying CO_2 fluxes. Limitations associated with the NMC method prompted the investigation of a new flexible approach for parameterizing the background error statistics that is more suited to atmospheric CO_2-DA.

Using the difference between CO_2 concentrations of two different models as a proxy for the background errors, spatial statistical tools are used to generate the background error statistics.

The resultant error correlation functions are consistent with the large scale structures in the background error. The new error statistics are also found to imply that errors are correlated over longer distances, indicating that the information from the assimilated observations can be used to reduce errors over larger regions. Experiments using GOSAT observation and subsequent evaluation with independent CO_2 observations demonstrated that accounting for the spatial and temporal variability in the fluxes is necessary to generate reliable and consistent atmospheric CO_2 concentrations. Because the CO_2 problem is dominated by the surface fluxes, the errors in their prescribed or even prior (in case they form part of the minimization control vector) values is of utmost importance. This study illustrates the improvement that can be made and therefore makes the clear case for including a good estimate of surface flux errors and error correlations in any method that is used for estimating the background errors of an atmospheric CO_2 data assimilation system.

The experiments also demonstrated a couple of caveats with the proposed approach. First, a judicious selection of the models is necessary to capture the maximum possible uncertainty in the background. Ideally one should attempt to generate the background error from an ensemble of models that would be more representative of the background error. Secondly, generation of the Δ-statistics requires separate analysis and computational time as compared to a forecast-based technique as the NMC method. The overall benefits from the new approach, however, outweigh these drawbacks in terms of defining a more physically consistent CO_2 data assimilation system. With the increasing interest in atmospheric CO_2 data assimilation systems in order to make efficient use of remote-sensing observations, the results from this work are expected to contribute to an overall improvement in data assimilation systems for carbon science applications.

Chapter 5

Inter-comparison of ensemble and variational data assimilation in the context of a CO_2 flux estimation problem

5.1 INTRODUCTION

Data Assimilation (DA) is best known as a tool in Numerical Weather Prediction (NWP; e.g. *Swinbank* [2010]) and has been applied to analyze complex datasets and estimate parameters in a variety of fields, including atmospheric constituent (e.g. *Lahoz and Errera* [2010]; *Elbern et al*. [2010]), oceanographic (e.g. *Haines* [2010]), and land surface (e.g. *Reichle* [2008]; *Houser et al*. [2010]) assimilation problems. In all applications, a DA system aims to optimally combine the information from the observations as well as a prior model estimate (or the background based on a model forecast), based on their respective uncertainty estimates.

DA methods for estimating CO_2 fluxes, similarly aim to constrain the spatial and temporal distributions of CO_2 sources and sinks by integrating atmospheric, terrestrial and oceanic data together into a common analysis framework. CO_2-DA applications (e.g. ensemble filter based methods - *Peters et al*. [2005]; *Feng et al*. [2009]; *Miyazaki et al*. [2011]; *Kang et al*. [2011]; *Chatterjee et al*. [in press], variational based methods - *Rayner et al*. [2005]; *Chevallier et al*. [2005b]; *Rödenbeck et al*. [2005]; *Baker et al*. [2006a], or hybrid approaches such as the Maximum Likelihood Ensemble Filter - *Zupanski et al*. [2007a]; *Lokupitiya et al*. [2008]) have been in vogue for nearly a decade and are viewed as an alternative to more traditional batch

inverse modeling schemes (Bayesian Inverse Modeling, e.g. *Enting* [2002]). Unlike DA, which uses a combination of numerical approximations and time-stepping approaches, the batch schemes solve the linear system of equations relating the fluxes and the atmospheric CO_2 observations in a single step. The DA approaches are attractive because of their computational efficiency (e.g. *Rayner* [2010]) but the exact impact of their underlying numerical approximations on the final estimates and their associated uncertainties is still ambiguous.

Recently, *Chatterjee et al*. [in press] pointed out that because of certain fundamental differences between the carbon flux estimation (i.e., the inverse framework) and the NWP/constituent (i.e., assimilation framework) problems: a) the performance of the DA methods are not necessarily equivalent for the two frameworks, and b) only under specific inversion scenarios are the DA methods able to perform optimally. Differences between the two frameworks are mainly driven by: a) the ill-conditioned and highly diffusive nature of the flux estimation problem, and b) the absence of an explicit dynamical model that can evolve a set of estimated fluxes forward in time. The lack of a dynamical model represents a loss of valuable information to the DA system. By propagating the state vector between different assimilation time steps, the dynamical model directly contributes to the growth of the eigenvalue spectrum of the state covariance matrix in certain preferred directions and decay in others (*Bengtsson et al*. [2003]; *Furrer and Bengtsson* [2007]). But for the CO_2 inverse problem, the absence of such a model coupled with the availability of sparse observational datasets may result in the DA approaches performing sub-optimally.

The authors are not aware of any study, specifically related to the CO_2 flux estimation problem, which attempts to evaluate the performance of different data assimilation techniques. This is unlike the weather community, where several studies have evaluated the strengths and

weaknesses of ensemble and variational approaches for different weather-related applications ranging from simple to chaotic non-linear systems (e.g. *Lorenc* [2003]; *Caya et al*. [2005]; *Fertig et al*. [2007]; *Kalnay et al*. [2007]; *Liu et al*. [2008b]; *Whitaker et al*. [2009]; *Buehner et al*. [2010a, 2010b]; *Jardak et al*. [2010]; *Zhang et al.* [2011]; also see the special collection of papers on inter-comparison at http://journals.ametsoc.org/page/Ensemble_Kalman_Filter). Apart from NWP-related comparison studies, DA methods have also been inter-compared for chemical (e.g. *Carmichael et al*. [2008]) and constituent (e.g. ozone – *Wu et al*. [2008]) assimilation problems. These comparison studies cannot be used as a baseline, however, because of differences amongst the flux estimation and the NWP/constituent data assimilation frameworks, as stated above.

The main purpose of this work is to fill this gap and build on the existing body of inter-comparison studies from the perspective of the CO_2 flux estimation problem. Specifically, this study aims to answer the following two questions: a) what is the relative performance of two state-of-the art DA approaches (Ensemble Square Root Filter – *Whitaker and Hamill* [2002] and 4-dimensional variational – e.g. *Talagrand* [2010]) for solving the CO_2 inverse problem?, and b) in the absence of a dynamical model for the inverse problem, can the DA approaches reproduce the flux estimates from a batch inverse modeling scheme?

To facilitate the inter-comparison, we consider here a one-dimensional passive tracer transport problem. The one-dimensional problem allows us flexibility to set up the inverse problem because multiple experiments can be simulated in a computationally efficient way. The low computational cost associated with the 1D problem enables the implementation of a batch least-squares method alongside the DA approaches. The DA estimates are compared to both the true signal and the batch estimates in order to isolate the degradation due to the underlying

numerical approximations. The final estimate in any DA technique is based on making several compromises between the computational cost, statistical optimality and physical realism of the assimilation problem. This study will assess whether these compromises allow the examined DA methods to be suitable long-term replacements of the batch technique, under different inversion conditions.

When designing the 1D problem, we focus on setting up a framework that allows us to examine an *under-determined and fine-scale* flux estimation problem. This set up is necessary to mimic the challenges of a true CO_2 flux estimation problem in which atmospheric mixing coupled with the sparseness of observations leaves the problem highly under-determined and ill-posed. The under-determined nature of the problem is accentuated by the need for estimating CO_2 fluxes at fine spatial and temporal scales. This is necessary to avoid aggregation errors (e.g. *Kaminski et al*. [2001]; *Michalak et al*. [2004]; *Gourdji et al*. [2012]) and improve our understanding of the fine-scale processes driving the carbon cycle. This shift in paradigm has primarily brought about a computational bottleneck in solving the batch inverse problem, as it requires the atmospheric transport model to be run either once per estimated flux region/period combination, or once per observation if an adjoint to the transport model is available. This has prompted the use of computationally efficient alternatives, such as DA methods in which the number of atmospheric transport model runs is proportional to the number of ensembles (in the ensemble approach) or the number of descent iterations (in the variational approach) rather than the number of estimated parameters or available observations.

Analogous to a real CO_2 flux estimation problem, no dynamical model is specified for solving the 1D problem using the DA methods. The experiments are specifically targeted to evaluate the impact of three factors on the two DA approaches: (a) impact of the observational

density and homogeneity, (b) impact of the model-data mismatch covariance, and (c) impact of the operational parameters of the DA system (i.e., ensemble size, number of iterations). Given the absence of a dynamical model, it is expected that the DA methods are highly sensitive to the information flowing in from the observations. Assessing the impact of the density and heterogeneity of the observation network and the quality of the observations is therefore necessary to gauge the optimality of the DA methods under different inversion scenarios. While determining the impact of the observational density and homogeneity, issues of sampling and convergence error are minimized by specifying a large number of ensemble members and descent iterations, for the EnSRF and the 4D-VAR[1], respectively. Operational constraints are subsequently imposed in the final set of experiments in order to investigate the interplay between the underlying numerical approximations and the influence of the observational density and homogeneity. Thus, the inter-comparison presented in this study not only addresses the fundamental differences between the DA methods but also the effect of the compromises necessary to make the algorithms practically feasible. The study is a first comparison of the EnSRF and the 4D-VAR for a simplified flux estimation problem, and can guide the development of future inter-comparison experiments with real data.

[1] *Typically in the DA community the term 4D-VAR is used to represent the 3-dimensional space plus time. In this study, the variational approach is applied to a 1-dimensional space plus time, which may suggest that the term '1+1D-VAR' may be more appropriate. Within the geophysical community, however, the term 1D in 1D-VAR specifically refers to the vertical column, and is quite popular for radio occultation data (e.g. Eyre et al. [1993]; Poli et al. [2002]), total column water vapor (e.g. Marécal and Mahfouf [2002]; Bauer et al. [2006]), cloud (e.g. Janiskova et al. [2012]) assimilation etc. Since in the current study, 1D refers to a single dimension along the horizontal space and not necessarily the vertical column, we persist with usage of the term 4D-VAR rather than the term 1D-VAR.*

5.2 EXPERIMENTAL FRAMEWORK

5.2.1 Estimation methods

Following *Enting* [2002], in a Bayesian framework, all the information that is used to constrain a set of parameters that are being solved for, can be represented by probability density functions (*pdf*s). If the *pdf*s can be approximated as Gaussian, then maximizing the posterior probability of the state is equivalent to finding the minimum of a quadratic objective function as shown in Equation 5.1:

$$J(\mathbf{s}) = \frac{1}{2}\left[\mathbf{z} - h(\mathbf{s})\right]^T \mathbf{R}^{-1}\left[\mathbf{z} - h(\mathbf{s})\right]^T + \frac{1}{2}\left[\mathbf{s} - \mathbf{s}^b\right]^T \left(\mathbf{Q}^b\right)^{-1}\left[\mathbf{s} - \mathbf{s}^b\right]^T \qquad 5.1$$

where **s** is a $m \times 1$ vector of the discretized state (for e.g., CO_2 flux), **z** is the $n \times 1$ vector of observations (for e.g., CO_2 observations), h is a forward model that is often a combination of an atmospheric transport model and an observation operator, **R** is the $n \times n$ model-data mismatch covariance, $\mathbf{s}^b$ is the $m \times 1$ prior estimate of the state, $\mathbf{Q}^b$ is the $m \times m$ error covariance of the prior estimate $\mathbf{s}^b$, and the superscript T denotes the matrix transpose operation. Note that in the case of atmospheric CO_2 inverse problem, h is linear and represented via its matrix form **H** (a.k.a. sensitivity matrix with dimensions $n \times m$, or Jacobian matrix). Typically, this matrix form of **H** represents the sensitivity of the observations **z** to the fluxes **s** (i.e. $H_{i,j} = \partial z_i / \partial s_j$).

The inverse problem as formulated via Equation 5.1 determines a least squares fit of the prior state estimate to the observations, subject to the constraints provided by the error covariance matrices. The estimate of **s** can be obtained via a batch Bayesian Inverse Modeling (BIM) approach (e.g. *Enting* [2002]). The analytical solution for the *a posteriori* estimate and the covariances are given by:

$$\hat{\mathbf{s}} = \mathbf{s}^b + \mathbf{K}\left(\mathbf{z} - \mathbf{H}\mathbf{s}^b\right) \quad 5.2$$

$$\mathbf{Q}^a = \left(\mathbf{I} - \mathbf{K}\mathbf{H}\right)\mathbf{Q}^b \quad 5.3$$

$$\mathbf{K} = \mathbf{Q}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{Q}^b\mathbf{H}^T + \mathbf{R}\right)^{-1} \quad 5.4$$

where $\hat{\mathbf{s}}$ is the posterior best estimate of the state and $\mathbf{Q}^a$ is the *a posteriori* covariance of the recovered best estimate. The diagonal elements of $\mathbf{Q}^a$ represent the predicted error variance ($\sigma_{\hat{\mathbf{s}}}^2$) of individual elements in $\hat{\mathbf{s}}$. As stated earlier in Section 5.1, for CO_2 inversion studies, generation of the matrix **H** requires an atmospheric transport model to be run either once per estimated state, or once per observation. The large number of forward model runs ultimately makes the batch approaches, such as BIM, computationally intractable for solving large-scale problems.

In the variational approach, Equation 5.1 is recast as a minimization problem and the solution $\hat{\mathbf{s}}$ is sought iteratively by an optimization algorithm. The overall approximation lies in the fact that only a small number of iterations are performed. The minimization can be stopped by artificially limiting the number of iterations or by requiring that the norm of the gradient decreases by a predefined amount during the minimization. Most minimization schemes (e.g., L-BFGS, *Nocedal and Wright* [2006]) rely on the availability of the gradient of the objective function ($\nabla J(\hat{\mathbf{s}})$) with respect to the state (or control vector in 4D-VAR terms).

$$\nabla J(\mathbf{s}) = \left(\mathbf{Q}^b\right)^{-1}\left[\mathbf{s} - \mathbf{s}^b\right] + \mathbf{H}^T\mathbf{R}^{-1}\left[\mathbf{z} - h(\mathbf{s})\right] \quad 5.5$$

Instead of analytically calculating the gradient, the adjoint of the forward transport model is used to compute the term $\mathbf{H}^T\mathbf{R}^{-1}\left[\mathbf{z} - h(\mathbf{s})\right]$, which is then added to $\left(\mathbf{Q}^b\right)^{-1}\left[\mathbf{s} - \mathbf{s}^b\right]$ in Equation 5.5 above. In this study, the 4D-VAR is implemented with successive overlapping windows to account for the long residence times of CO_2 in the atmosphere. If a single long window is

employed for a real CO_2 flux estimation problem, the computational cost associated with calculating the inverse of the prior covariance matrix $\mathbf{Q}^b$ in Equation 5.5 increases significantly. Alongside the CO_2 flux errors are both spatially and temporally correlated (e.g. *Chevallier et al*. [2012]), which implies that the prior error covariance matrix has off-diagonal terms. Any algebraic manipulations, such as taking inverses or calculating the square roots for preconditioning purposes, thus become operationally cumbersome.

In the absence of a dynamical model, we find the 4D-VAR implementation to be similar to the FGAT-3DVAR (*Massart et al*. [2010]) variant occasionally used within the NWP community. The main caveat with the variational approach is that a direct estimate of the analysis error is not available (no clear analogue of Equation 5.3). Mathematically this can be obtained from the inverse of the Hessian (e.g. *Le Dimet et al*. [2002]; *Rödenbeck* [2005]; *Meirink et al*. [2008]). But computational challenges restrict the calculation and storage of the Hessian for large scale problem. Recent studies have shown that computationally efficient alternatives do exist (e.g. *Shutyaev et al*. [2009]; *Cheng et al.* [2010]), which may have potential applicability for the CO_2 flux problem but have not been pursued in this study.

In the ensemble filter approach, the key innovation is to work in a reduced subspace of the error covariance matrices. Observations are assimilated to update the ensemble representation of the error covariance matrices. The optimal analysis states and an estimate of the analyses error are determined in a similar fashion to Equations 5.2 and 5.3 but the calculation of $\mathbf{Q}^b\mathbf{H}^T$ and $\mathbf{H}\mathbf{Q}^b\mathbf{H}^T$ are approximated by running the transport model with the ensemble members directly:

$$\mathbf{H}\mathbf{Q}^b\mathbf{H}^{\mathrm{T}} \approx \frac{1}{N-1}\left[h(\mathbf{s}')\right]\left[h(\mathbf{s}')\right]^T \qquad 5.6$$

$$\mathbf{Q}^b\mathbf{H}^{\mathbf{T}} \approx \frac{1}{N-1}(\mathbf{s}')\left[h(\mathbf{s}')\right]^T \quad 5.7$$

While several variants of the ensemble approach are available, here we have used a serial ensemble square root filter (*Whitaker and Hamill* [2002]), which is implemented in a fixed-lag smoother form (e.g. *Whitaker and Compo* [2002]; *Chatterjee et al*. [in press]). The localization scheme (i.e., to cut down spurious noise in the ensemble) is based on *Houtekamer and Mitchell* [2001] using a fifth-order Gaspari-Cohn function (*Gaspari and Cohn* [1999]) while the adaptive inflation algorithm (i.e., to counter spurious variance deficiency among the ensemble members) is based on *Anderson* [2009]. Implementation of these algorithms within an ensemble smoother framework is described in further detail in *Chatterjee et al*. [in press] (also see Chapter 6 - Section 6.2). Inspite of these additional algorithms, the overall implementation of EnSRF is surprisingly simple and computationally efficient.

The setup of 4D-VAR (overlapping time windows) and EnSRF (fixed-lag smoother) reflect state of the art implementations of these two DA approaches, keeping in mind the nature of the atmospheric CO_2 process. We tested other flavors of the ensemble filter (EnKF with perturbed observations) and the variational approach (PSAS) but found that the overall conclusions from the inversion experiments remain largely consistent across these algorithmic choices. We encourage the reader to look at *Lorenc* [2003] and *Nichols* [2010] (and references therein) for a more mathematical and detailed exposition on these DA methods.

5.2.2 Problem description

A 1-dimensional (1D) advection-diffusion problem of a passive tracer is selected to emulate the CO_2 flux estimation problem. In the 1D illustration, the passive tracer represents atmospheric CO_2. Tracer fluxes get released from a series of locations over a finite duration and

get transported by a tracer transport model that encapsulates the physics of advection and diffusion. No sink is specified in the model as a result of which there is a gradual buildup of the passive tracer within the domain. Subsequently, observations of the tracer are obtained at various locations and times within the domain. The locations and times of the observations as well as their precision can be regulated to simulate a variety of inversion scenarios. The inverse problem constitutes of using the noisy tracer observations along with the transport information to infer the original tracer fluxes.

In the following description, the units of mass, length and time are reported as [M], [L] and [T] to keep the problem generic. Both the length of the 1D domain and the time period of the experiment are arbitrarily discretized. The parameters for the experiment are: the grid size Δx = 1 [L]; the domain length x = 300 [L]; the time step of release Δt = 1 [T]; the total number of time periods over which the tracer flux is released is 35; the longitudinal dispersion coefficient D_L = 2.0 [L^2/T]; and the advection velocity v = 50.0 [L/T].

The tracer flux **s** [$ML^{-1}T^{-1}$] that is released (Figure 5.1A) is modeled as:

$$\mathbf{s}\left(\mathbf{x}_r,\mathbf{t}_r\right) = 0.25\left(36-\mathbf{t}_r\right)\exp\left[-\frac{\left(\mathbf{x}_r-70\right)^2}{200}\right] + \exp\left[-\frac{\left(\mathbf{x}_r-130\right)^2}{50}\right] + \exp\left[-\frac{\left(\mathbf{x}_r-150\right)^2}{50}\right] + 0.25\left(\mathbf{t}_r\right)\exp\left[-\frac{\left(\mathbf{x}_r-220\right)^2}{200}\right] \quad 5.8$$

where $\mathbf{x}_r$ represents the locations along the 1D domain ($\mathbf{x}_r$ = 1, 2, 3…299, 300 [L]) over which the tracer flux is released at times $\mathbf{t}_r$ ($\mathbf{t}_r$ = 1, 2, 3…34, 35 [T]). Equation 5.8 is designed to model two large peaks with fluctuating amplitudes (Figure 5.1B) between 50-100 [L] and 200-250 [L], and a smaller consistent double peak (Figure 5.1B) between 100-150 [L]. Even though the spatial tracer flux profiles are different for each time period, the spatially averaged flux has a constant value of 0.84 [$ML^{-1}T^{-1}$] across all time periods. Note that the true tracer flux **s** is used only for simulating the plume but then considered unknown during the analysis.

The tracer is sampled at the locations $\mathbf{x}_o$ ($\mathbf{x}_o$ = 1, 2, 3,…, 299, 300 [L]) for 35 consecutive time periods ($\mathbf{t}_o$= 1.5, 2.5, 3.5…34.5, 35.5 [T]) to obtain the observational dataset $\mathbf{z}$ [ML^{-1}], such that the observations times are offset from the release times. Initial random error with a pre-specified variance (10 [M^2L^{-2}]) is added to the tracer observations to simulate measurement, transport, aggregation, and representation errors. Later in the study, different configurations of $\mathbf{x}_o$ and a higher error variance ($\sigma_{\mathbf{R}}^2$) are prescribed to test the impact of these parameters.

The tracer observations ($\mathbf{z}$) and the tracer fluxes ($\mathbf{s}$) are related in the following fashion:

$$\mathbf{z} = \mathbf{H}\mathbf{s} + \nu, \quad \text{where } \nu \sim N(0, \mathbf{R}) \tag{5.9}$$

where $\mathbf{H}$ is the sensitivity matrix that is generated using a 1D tracer transport model as:

$$H_{o,r} = \frac{1}{2}\left[erfc\left(\frac{(x_o - x_r) - vt_o}{2\sqrt{D_L t_o}} \right) - erfc\left(\frac{(x_o - x_r) - v(t_o - t_r)}{2\sqrt{D_L (t_o - t_r)}} \right) \right] \tag{5.10}$$

where $\mathbf{x}_r$ and $\mathbf{t}_r$ are the tracer flux release locations and times, $\mathbf{x_o}$ and $\mathbf{t}_o$ are the tracer observation locations and times, *erfc* represent the complimentary error function. The tracer transport model

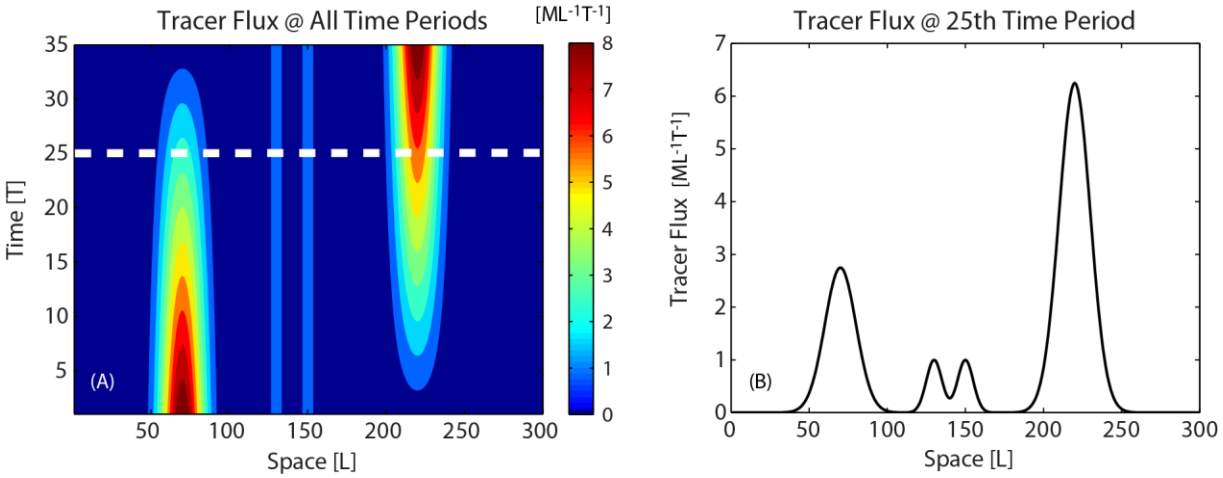


Figure 5.1 *– (A) Filled contour plot of the tracer flux, over the entire domain and for different time periods. (B) Flux profile for a particular time period corresponding to the dashed white-line in (A).*

embedded in Equation 5.10 assumes conservation of mass and is based on a well-known one-dimensional analytical solution for contaminant transport in the groundwater literature (e.g. *Ogata and Banks* [1961]; *Runkel* [1996]).

The tracer observations obtained at a particular time step are sensitive to the tracer flux released at multiple previous time steps. Given that the total length of the domain is 300 [L] and the advection velocity is 50 [LT^{-1}], the maximum residence time of the tracer within the domain can be 6 [T]. Based on the exponential construction of Equation 5.8, however, the fluxes start being released from approximately 50 [L] onwards, in which case their residence time in the domain is reduced to 5 [T]. This means that an observation taken at time $\mathbf{t}_o$ [T] provides information about the tracer flux approximately up to time $\mathbf{t}_o$ - 5 [T]. Hence, based on the physics driving the advection-diffusion problem, in all subsequent experiments the lag window size for the DA schemes is set to 5 [T].

The final piece of information necessary for setting up the inverse problem is the prior estimate ($\mathbf{s}^b$) of the tracer flux and its error characteristics. $\mathbf{s}^b$ is constant across all time periods:

$$\mathbf{s}^b\left(\mathbf{x}_r,\mathbf{t}_r\right) = \exp\left[-\frac{\left(\mathbf{x}_r - 150\right)^2}{2000}\right] \qquad 5.11$$

Its error covariance matrix $\mathbf{Q}^b$ is based on an exponential decay model, with a correlation length ($3l_{\mathbf{Q}}$) of 90 [L] and variance ($\sigma_{\mathbf{Q}}^2$) of 3 [$M^2L^{-2}T^{-2}$].

The 1D framework has been designed to capture most of the characteristic features of the CO_2 flux estimation problem. For a real CO_2 inversion, these units would typically have been [µmol/(m^2s)] for $\mathbf{s}$ and $\sigma_{\mathbf{Q}}$, [ppm] for $\mathbf{z}$ and $\sigma_{\mathbf{R}}$, [ppm/ µmol/(m^2s)] for h_{1D} and [km] for $l_{\mathbf{Q}}$.

5.2.3 Experiments

Experiments are designed to explore the impact of three factors on the ability of the DA methods to solve the inverse problem: (a) the observational density and homogeneity, (b) the model-data mismatch covariance, and (c) the computational constraints of the DA system (i.e., ensemble size, number of iterations). In all the experiments, the size of the state vector or the total number of fluxes to be inferred is 10500 × 1 (i.e. [300*35] × 1).

The first set of experiments (Table 5.1 – Experiments A through C) aims to investigate the effect of the density and homogeneity in space and time of the observational network. Three different observational networks are designed (Figure 5.2). In the first network configuration (denoted as REF – 'reference observational set', same as outlined in Section 5.2.2), observations are obtained at all locations ($\mathbf{x}_o$ = 1, 2…300 [L]) and for all 35 time periods ($\mathbf{t}_o$= 1.5, 2.5, 3.5 …34.5, 35.5 [T]). The total number of observations available is 10500 (300*35). In the second network configuration (denoted as HM – 'homogeneous'), observations are obtained at

Table 5.1 – *Summary of the experiments outlined in Section 5.2.3. The following parameters are held constant for all the experiments in this study:* $\boldsymbol{s}^b$ *(Equation 5.11),* $3l_Q = 90$ *[L] and* $\sigma_Q^2 = 3$ *[M²]*

Impact Considered	**Experiment**	**Observation Parameters**		**DA Parameters**	
		Network	Variance σ_R^2	Ensemble Size	Descent Iterations
Observational density and homogeneity	A	REF	10	1000	250
	B	HM	10	1000	250
	C	HT	10	1000	250
Model-data mismatch	AR	REF	400	1000	250
	BR	HM	400	1000	250
	CR	HT	400	1000	250
Operational limitations	AO	REF	10	100	25
	BO	HM	10	100	25
	CO	HT	10	100	25

25 equally spaced locations within the 1D domain ($\mathbf{x}_o$ = 10, 22, 34…298 [L]), which are maintained for all 35 time periods ($\mathbf{t}_o$= 1.5, 2.5, 3.5…34.5, 35.5 [T]). The total number of observations is thus reduced to 875 (25*35). In the final configuration (denoted as HT - 'heterogeneous'), observations are taken at 25 random locations, which vary arbitrarily over the 35 time periods ($\mathbf{t}_o$= 1.5, 2.5, 3.5…34.5, 35.5 [T]). Similar to HM, in HT the total number of observations remains 875 (25*35) but the observations are neither uniform in space nor consistent over time. Note that unlike REF, both HM and HT represent under-determined inversion problems where the total number of observations is significantly lower than the number of states to be estimated. In reality the HT network configuration scheme is the closest to current CO_2 monitoring networks where different monitoring locations (ground-based or remote-sensing) can come online and go offline over different periods. For all the three network configurations, random error with a variance of 10 [M^2L^{-2}] is added to the observations to make the problem realistic. This allows us to specify the model data mismatch covariance matrix **R** (in Equation 5.1) with its diagonal values corresponding to variance of the errors introduced in the

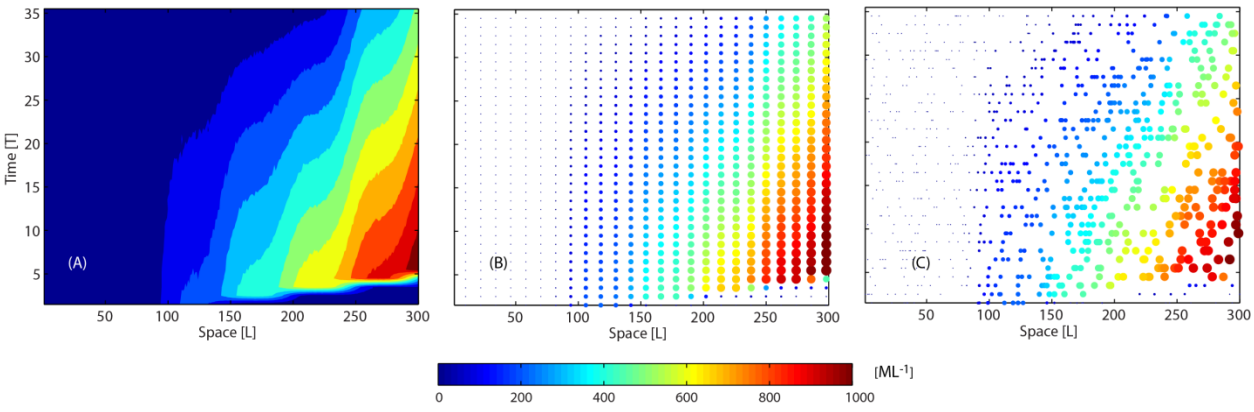


***Figure 5.2** – Observations of the tracer obtained from the three network configurations - REF (A), HM (B) and HT (C). Note that – (i) going from the REF to the HM and the HT networks, the total number of observations decreases, and (ii) going from the HM to the HT network, the observational fields become more heterogeneous in space and time.*

synthetic observations. All the observations are assumed to have the same model-data mismatch error without any temporal correlation.

The second set of experiments (Table 5.1 – Experiments AR through CR) examines the effect of the model-data mismatch covariance on the best estimates and their associated uncertainties. The model data mismatch covariance matrix captures not only the errors in the observations but also errors due to other sources, such as the transport model, representation and aggregation errors (e.g. *Engelen et al*. [2002]). For all the network configurations, the variance of the random errors is increased to 400 [M^2L^{-2}] and the diagonal values of the model data mismatch covariance matrix **R** increased accordingly. All the other parameters are maintained the same as the first set of experiments.

The third set of experiments (Table 5.1 – Experiments AO, BO and CO) explores the impact of the operational constraints in implementing any DA system. To minimize the numerical approximations and avoid any form of sampling or convergence error in the first two set of experiments (Table 5.1 – Experiments A-C and Experiments AR-CR), the ensemble size (for the EnSRF) and the number of descent iterations (for the 4D-VAR) are set to 1000 and 250, respectively. The number of descent iterations are prescribed to be lower than the number of ensemble members, keeping in mind that 4D-VAR typically requires more model integrations (i.e., both forward and adjoint model run) than EnSRF. Given that it is not feasible to either run a large number of ensemble members or specify a large number of descent iterations for any real atmospheric application, these are reduced to 100 ensemble members for the EnSRF and 25 descent iterations for the 4D-VAR. The noise added to the observations is reduced to 10 [M^2L^{-2}], however, so that a direct comparison can be made with Experiments A-C.

5.3 RESULTS

Results from the nine experiments are interpreted: (a) at native and aggregated spatial scales, and (b) in terms of the individual performance of EnSRF and 4D-VAR relative to the truth and the BIM estimates. Three metrics - the root mean square difference (RMSD), the correlation coefficient (CC) and the standard deviation (SD) are calculated between the flux estimates and the truth at the grid-scale (Figure 5.4) or at spatially aggregated scales (Figure 5.5). All metrics are averaged across 30 time periods ($\mathbf{t}_r$ = 6, 7, 8…34, 35 [T]) after discarding the first 5 time periods as spin-up.

5.3.1 Impact of observational density and homogeneity

Figure 5.3A-C shows the recovered tracer fluxes for the three network configurations. For the REF-network in Figure 5.3A, all three analyses methods perform extremely well in: a) recovering the true flux that was released during this time period (Panel A), and b) fitting the observations within the specified errors (results not shown). Both the 4D-VAR and the EnSRF estimates qualitatively capture the flux profile, including its large and the small scale peaks. The estimates obtained from the three analyses methods are not specific to the time period shown in Figure 5.3 but remain consistent across other time periods. As indicated in Figure 5.4A, all the three methods have a high CC (~0.97), low RMSD (~0.3 [$ML^{-1}T^{-1}$]) and identical standard deviation (~1.5 [$ML^{-1}T^{-1}$]) as the true fluxes. The performance of all three methods degrades as the observation density and homogeneity changes in going from Experiments A to C. This is evident by looking at Figure 5.3B-C where the methods fail to capture precisely the smaller double peak around 100-200 [L] while the Taylor diagram in Figures 5.4B-C captures the resultant drop in CC and a corresponding increase in RMSD.

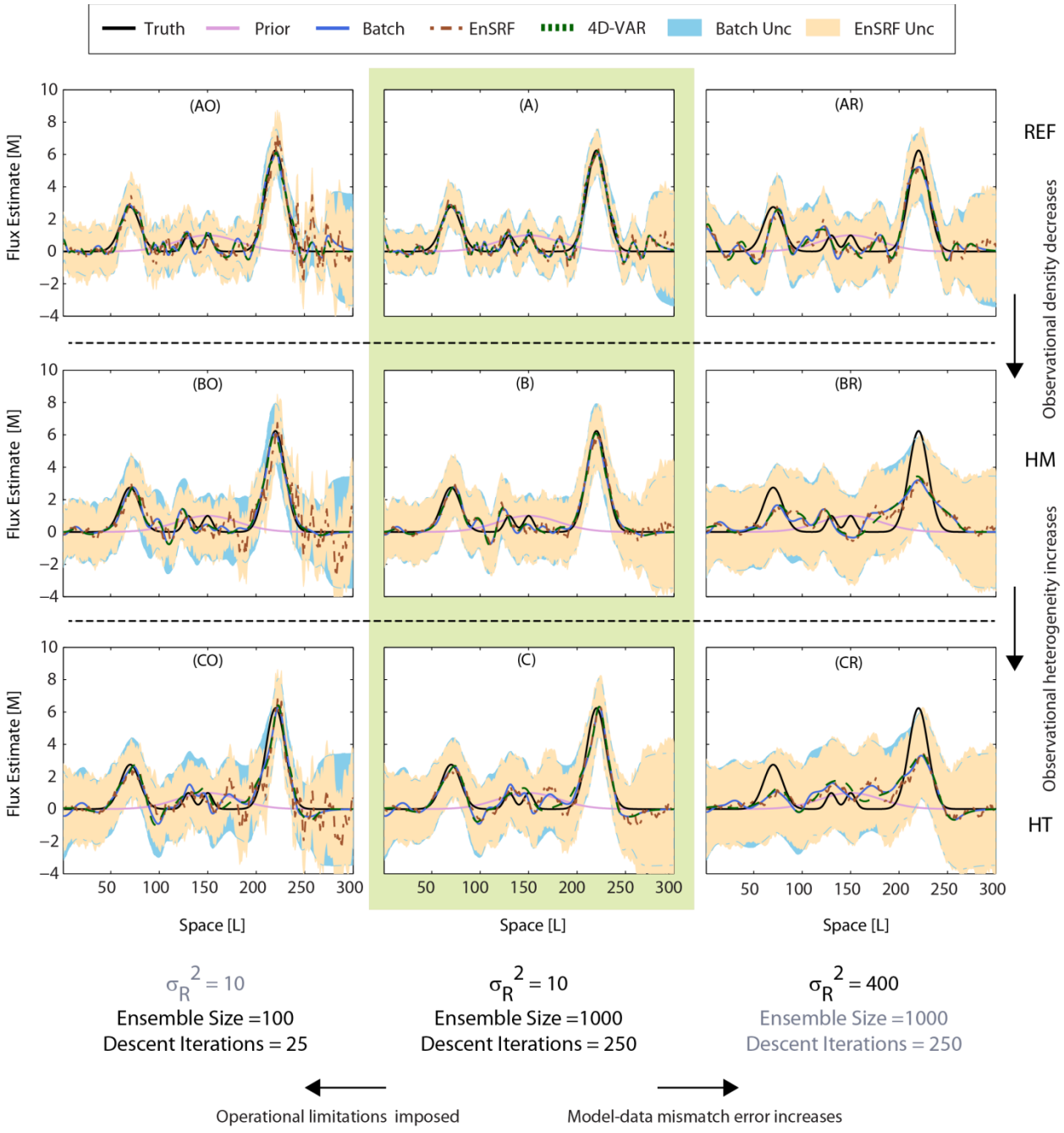


Figure 5.3 *– True and estimated tracer fluxes and associated uncertainties for the different experiments in this study. All values are shown for the 25th time period, which is assumed to be representative of other time periods. The panel numbers correspond to the different experiments outlined in Table 5.1.*

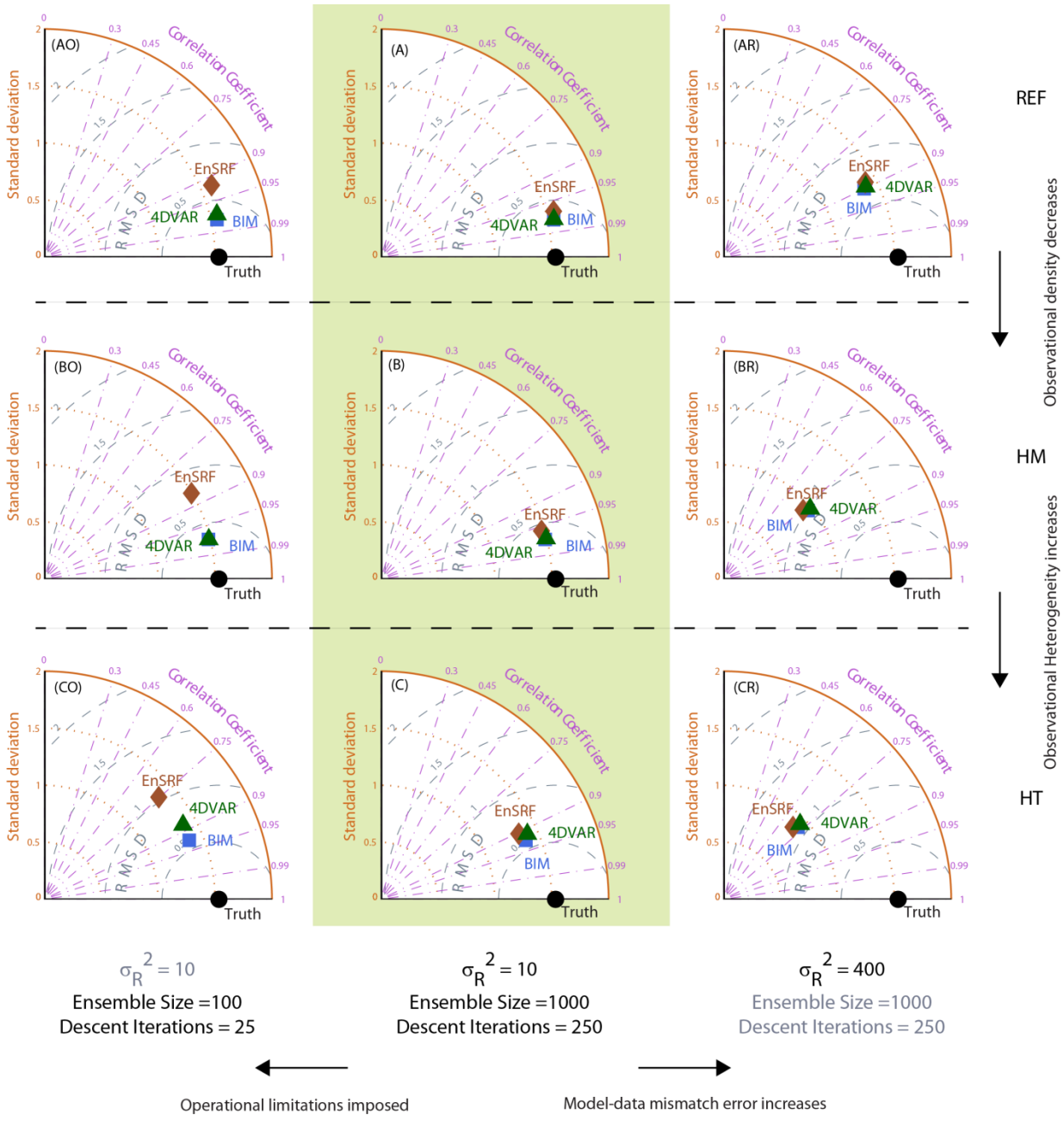


Figure 5.4*- Performance of the BIM, the EnSRF and the 4D-VAR for the different experiments outlined in Table 5.1. For each experiment, statistics are calculated between the estimates and the true fluxes across all locations and all 30 time periods and represented on a Taylor diagram (Taylor [2001]). For each Taylor diagram the true flux is represented by a point from the origin along the abscissa ('Truth'). All other points ('BIM', 'EnSRF', '4D-VAR') that represent the estimated fluxes, are positioned such that - their standard deviation is the radial distance from the origin, the correlation coefficient between the estimates and the truth is the cosine of the azimuthal angle, and the root mean square difference (RMSD) between the estimates and the truth is the distance to the observed point. In the limit of perfect agreement, RMSD should approach zero, CC should approach unity, and the SD of the estimates should be the same as the truth.*

In general, for observations with spatially uncorrelated error, decreasing the observation density is expected to decrease the analysis accuracy. The response of the two DA approaches mirror the BIM estimate in such cases, including the inference of a wrong flux pattern for the HT network around 100 [L] in Figure 5.3C. This indicates that in the absence of any operational limitations, if all other parameters of the inverse problem are the same, then the DA estimates are consistent with the BIM estimate. Neither the lack of a dynamical model nor the under-determined nature of the inverse problem impedes the ability of the DA methods in performing as optimally as the BIM approach.

Despite the lower quality of the best estimates, both the BIM and the EnSRF are able to capture the true flux profile within their 95% confidence bounds. The EnSRF uncertainty estimates are initially consistent with the BIM uncertainty estimates for the REF network (Figure 5.3A). In fact, the ratio of the predicted standard deviation of the individual flux estimates in EnSRF ($\sigma_{\hat{s}_{EnSRF}}$) to those from the BIM ($\sigma_{\hat{s}_{BIM}}$) is approximately 0.98, i.e. on average EnSRF underestimates the a posteriori uncertainties by 2% relative to BIM. Going from the REF to the HT scheme, the EnSRF over-estimates the uncertainty by 2% and 6% for HM and HT, respectively. For example, for the HT network (Figure 5.3C) the EnSRF over-estimates the error towards the end of the 1D domain (250 – 300 [L]).

In the EnSRF framework, the uncertainties are directly related to the ensemble spread. In the absence of a dynamical model, there is little source of variability for the ensemble to maintain a consistent spread. As observations are assimilated the ensemble tends to collapse to the ensemble mean and the adaptive inflation has to compensate for this degeneracy by inflating the ensemble spread. In the HT case, however, the inflation technique has a delayed response in adjusting to the heterogeneity in the observation network, as different observation locations

come into and out of the network. For the adaptive inflation component to work properly, we find it is beneficial to have a consistent set of observations to maintain a reasonable ensemble spread. It is worthwhile to mention here that the magnitudes of the inflation factors are very small in Experiments A-C. This is not surprising given that a large number of ensemble members have been specified and hence, the sampling error is quite low. Even without the application of inflation and localization (results not shown here), the EnSRF estimates have a high CC (~0.93), low RMSD (~0.66 [$ML^{-1}T^{-1}$]) but slightly higher SD (~1.78 [$ML^{-1}T^{-1}$]) to the true fluxes (SD = 1.58 [$ML^{-1}T^{-1}$]).

Overall, we find that both the 4D-VAR and the EnSRF performance reasonably match the BIM performance for all three observational network configurations. Even though small discrepancies are noticeable, the impact of a sparse and/or heterogeneous observational network is similar for DA as for the BIM approach.

5.3.2 Impact of model-data mismatch covariance

Panels AR-CR in Figures 5.3 and 5.4 (right hand column) illustrate the impact of the model-data mismatch covariance. For all the network configurations the quality of the estimates degrades when a higher model-data mismatch error is prescribed (comparing Panels A and AR, B and BR, C and CR) albeit the heterogeneous network estimates are the worst. Increase in $\sigma_{\mathbf{R}}^2$ also increases the uncertainty of the estimates, indicative of the decreased confidence in the analysis. Analogous to the first set of experiments, the DA approaches respond similarly to the BIM as the model-data mismatch covariance changes and both the 4D-VAR and the EnSRF track the BIM estimates quite well for all the three experiments. Figure 5.4AR-CR confirm that the estimates from all the three analyses methods have a lower CC, higher RMSD, and lower SD when compared to Panels A to C.

The standard deviation of the estimates changes considerably between Panels AR (~ 1.45-1.5 [$ML^{-1}T^{-1}$] for the three methods) and CR (~ 0.94-1.00 [$ML^{-1}T^{-1}$] for the three methods). Increasing the $\sigma_{\mathbf{R}}^2$ to 400 [M^2L^{-2}] results in the analysis rejecting the information from the observations and giving more weight to the prior, yielding overly smooth *a posteriori* estimates.. A typical example of this is seen by comparing the estimated peak around 50-100 [L] in Figure 5.3C and CR. Estimates in both these panels are based on the same observational fields but the estimates in Figure 5.3CR are unable to capture the amplitude of the peak in the true flux signal.

The inability to capture the true amplitude is slightly mitigated for the BIM and the EnSRF due the fact that these approaches infer higher uncertainty on the estimates. Both these estimates do capture the true flux profile within their respective 95% uncertainty bounds. As the observation network becomes sparser and more heterogeneous, the EnSRF slightly over-estimates the BIM uncertainty, by 3% (HM) and 5% (HT), while it underestimates the uncertainty by only 1% for the reference network. In spite of over-estimating the uncertainties, the EnSRF exhibits a clear advantage over the 4D-VAR by providing analyses errors as part of the estimation procedure.

Experiments AR-CR reconfirm that in the absence of operational limitations, increase (or decrease) in the model-data mismatch covariance do not make the DA methods perform any worse (or better) than the BIM approach.

5.3.3 Impact of operational constraints

Panels AO-CO in Figures 5.3 and 5.4 (left-hand column) shows the performance of the DA approaches when the number of ensemble members and the descent iterations are reduced. Reducing these parameters impacts the performance of the DA approaches negatively, and the impact is further intensified as the observational network becomes more heterogeneous. We find

that the 4D-VAR and the EnSRF performance are impacted differently as the numerical approximations come into play. These are discussed separately below.

For the 4D-VAR, an inadequate number of iterations fails to find the global minimum of the quadratic objective function (convergence results not shown here). When the observation fields are heterogeneous, the minimization has more difficulty in finding the closest path towards the true minimum. Thus, comparing Panels AO, BO and CO in Figure 5.4, the 4D-VAR estimates are worst for the HT network configuration. In general, we found that for the HT network, 4D-VAR needs approximately 150 iterations to converge completely as several iterations are wasted in zigzagging very slowly towards the optimum. Conversely, for the REF network 4D-VAR required less than 50 iterations to reach full convergence. When only 25 iterations are specified, expectedly the negative performance of the 4D-VAR is more visible for Experiment CO than Experiments AO and BO that have a relatively homogeneous network.

For all the three experiments, however, the magnitude of the posterior objective function is reduced relative to the prior, indicating that a more probable posterior state has been found compared to the prior. As pointed out by *Rödenbeck* [2005] in the initial iterations, the minimization determines the large-scale gradient while in subsequent iterations fine-scale tuning is performed to capture the optimum. By artificially limiting the number of iterations, the ability of 4D-VAR to make any small scale changes is hindered yet for Experiments AO and BO (Figure 5.4AO and BO), the 4D-VAR estimates are reliably close to the BIM estimates.

For the EnSRF, the degradation can be solely attributed to sampling error caused by a limited ensemble size. This reduces the estimation accuracy (both flux estimates and the uncertainties) and makes the filter sensitive to the observational density. As was stated earlier in Section 5.2.1, a Schur-based localization scheme has been implemented for the EnSRF. Since the

localization length scale is dependent on the ensemble size, when an ensemble size of 1000 is used (Experiments A-C, or AR-CR) a long localization length scale of 90-120 [L] is sufficient. The localization length scale is determined subjectively based on sensitivity tests, and hence a range of values (i.e., 90-120 [L]) are acceptable for which the EnSRF estimates are not contaminated by spurious noise.

Reducing the ensemble size to 100 necessitates a shorter localization length scale. Even with the application of localization, the spurious noise in the estimates overwhelms the information available from the observations. It is found beneficial to have two different length scales now, 10-30 [L] for the REF network and 45-60 [L] for the sparser networks. Specifying a longer localization length scale than 30 [L] for the REF network led to a divergence of the EnSRF system. In this case the spurious noise in the ensemble outweighs the positive impact of the observations. The complicated interplay between the ensemble size and the observational density makes it difficult to identify a mathematical or physical basis for selecting an appropriate localization length scale.

In terms of the recovered uncertainty, the sampling error plays a dominant role in determining the ability of the EnSRF to correctly represent the error in the ensemble mean. The EnSRF uncertainty estimates are close to the BIM uncertainty estimates but under/over-estimate at different estimation locations, rather randomly for the three network configurations, REF (over-estimation by 10%), HM (under-estimation by 9%) and HT (under-estimation by 3%). For Experiment CO, the EnSRF estimates fail to capture the true tracer flux within the 95% uncertainty bounds (see around 200-210 [L] in Figure 5.3CO).

An important caveat here is that the results for both the DA approaches could be potentially improved through further tuning of each algorithm. For example, implementation of

pre-conditioning algorithms to reach faster convergence, or more sophisticated localization schemes to dampen the spurious noise in the ensemble, may provide slightly different responses and reduce the error incurred due to the numerical approximations. In spite of having state-of-the-art covariance localization and adaptive inflation algorithms, once the underlying numerical approximations come into play: a) the EnSRF fails to reproduce the BIM estimate, with the EnSRF performance decreasing as the observation network becomes sparser and more heterogeneous, and b) the 4D-VAR fails to reproduce the BIM estimate when the observation network is heterogeneous but still performs better than the EnSRF. The better performance of the 4D-VAR is offset by the fact that the EnSRF provides explicit uncertainty bounds on the recovered flux estimates.

The DA approaches are specifically sensitive to the information flow from the observations because of the lack of a dynamical model. As discussed in Section 5.1 the dynamical model adds to the information content in the system. Identifying/developing an appropriate dynamical model relevant to the CO_2 flux estimation problem and subsequently repeating the experiments may allow an assessment of the interplay between the operational constraints and the observational network better. This task is beyond the scope of the current study but should be targeted in future studies to better understand the performance of DA approaches for the CO_2 inverse problem relative to applications within an assimilation framework.

5.3.4 Examining results at aggregated spatial and temporal scales

In the previous sections, the performance of the DA and the BIM approaches have been analyzed and reported at the actual estimation scales (both space and time). Typically in current

CO_2 studies, the fine-scale analyses are averaged spatially and/or temporally to coarser scales (for e.g., daily grid-scale fluxes averaged to monthly continental scales) and then examined.

When the estimates from the three analyses methods are aggregated across all space (i.e., across all locations $\mathbf{x}_r$ = 1, 2, 3…299, 300 [L]) and time (i.e. across all analyzed periods $\mathbf{t}_r$ = 6, 7, 8…34, 35 [T]), it is difficult to gauge any degradation in the EnSRF or the 4D-VAR. For example, after both spatial and temporal aggregation, the EnSRF and the 4D-VAR estimates recover the true flux and the BIM flux estimates precisely except for the two most challenging setups (Experiments CR and CO). Even in these cases the difference between the EnSRF and the 4DVAR with the BIM estimates is less than 5%. Likewise if the results are aggregated in time and examined spatially, the DA approaches are still able to recover the BIM estimates and the true fluxes.

Next the domain is divided into two segments, i.e. 1-150 [L] and 151-300 [L], and the true flux and the flux estimates are aggregated over each of these segments and examined across time. Since the true flux has fluctuating amplitudes between 50-100 [L] and 200-250 [L], it is possible to examine the ability of the inversion methods in capturing these distinct temporal variations. Quantitative metrics are calculated between the estimates and the true flux at the aggregated spatial scales and plotted on a Taylor diagram (Figure 5.5). For all the 9 experiments, the 4D-VAR is able to match the temporal variation of the spatially aggregated BIM estimates better than the EnSRF. The 4D-VAR iterative algorithm reliably captures the large scale patterns, as a result of which at any given time period and for any experiment, the 4D-VAR estimates are closest to the BIM. Even when the number of descent iterations are reduced (Panel AO-CO in Figure 5.5), at the spatially aggregated scales, the 4D-VAR estimates have negligible difference from the BIM. Comparing Panel CO in Figures 5.4 and 5.5 demonstrate that the

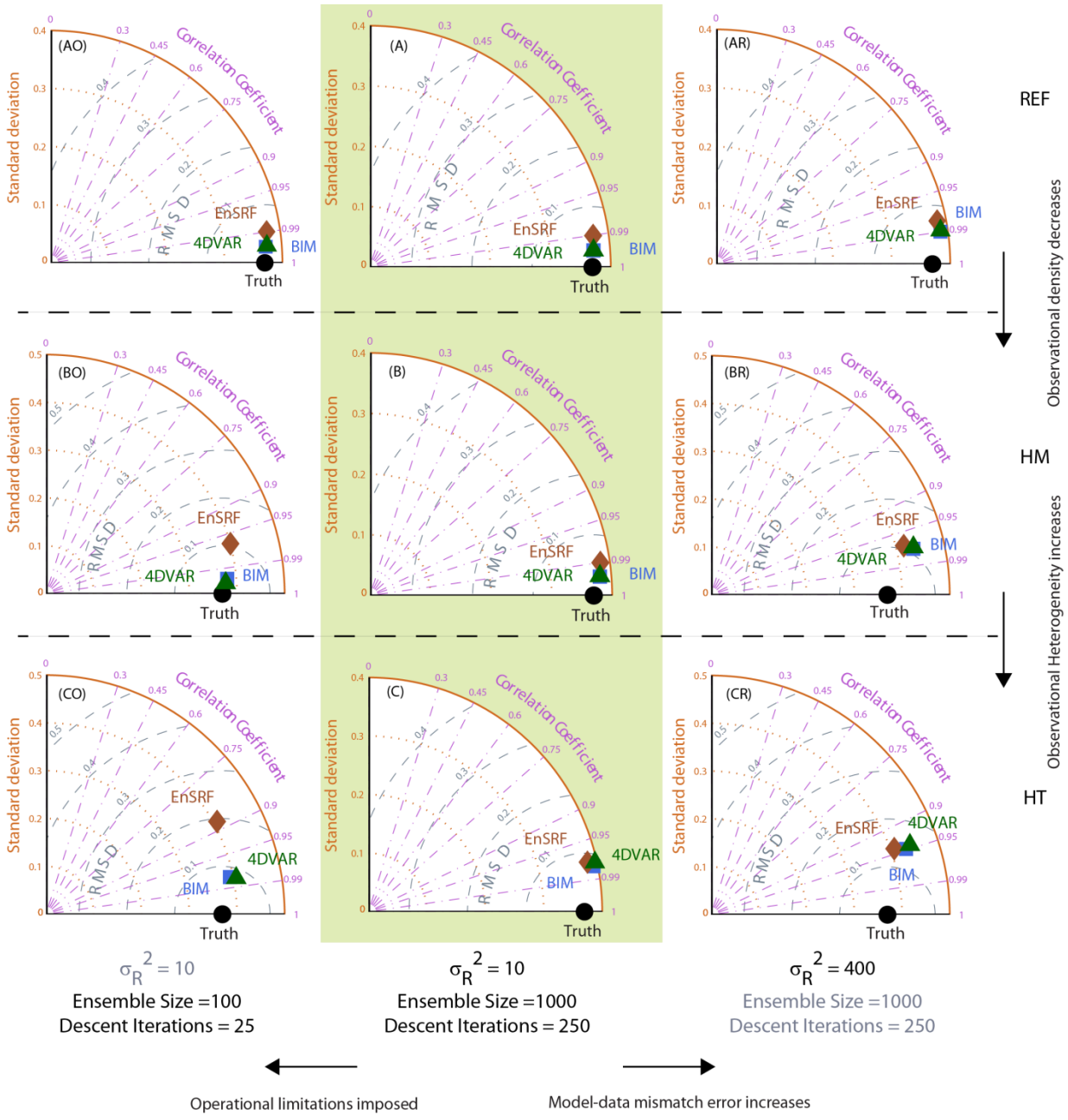


Figure 5.5 *- - Performance of the BIM, the EnSRF and the 4D-VAR at spatially aggregated scales for the different experiments outlined in Table 5.1.*

differences at fine scale are substantially reduced by aggregating the estimates to a coarser resolution.

Even though the EnSRF has been applied as a smoother, there is no dynamical model to evolve the information between assimilation time periods. For all the experiments that are run without any operational constraints, the EnSRF estimates are always contaminated with small sampling error. This gets washed out when the estimates are spatially aggregated but still EnSRF estimates are more erratic than the 4D-VAR estimates. When the number of ensemble members is reduced (Panel AO-CO in Figure 5.5), however, both the sampling error and the observational density and homogeneity starts playing a role. For example in Figure 5.5, Panels AO and CO, the CC between the spatially aggregated EnSRF and the true flux estimates drop from 0.99 to 0.90 and the RMSD increases from 0.03 to 0.21 [$ML^{-1}T^{-1}$]. The standard deviation of these spatially aggregated EnSRF estimates also increase from 0.37 to 0.40 [$ML^{-1}T^{-1}$] compared to the true flux, which has a standard deviation of 0.36 [$ML^{-1}T^{-1}$]. This indicates that even the spatially aggregated EnSRF estimates become more unpredictable as the observation network gets sparser and heterogeneous.

Analysis at aggregated scales demonstrates if operational constraints are not imposed, both the DA approaches provide aggregated estimates that are close to the aggregated BIM estimates. Even when operational constraints are taken into account, the aggregated flux estimates for the EnSRF and the 4D-VAR (Figure 5.5CO) have higher CC and lower RMSD than the corresponding values at the fine scale (Figure 5.4CO). This is encouraging from the perspective of a real CO_2 flux estimation problem, as it implies that the DA flux estimates can serve as reliable alternative for the BIM flux estimates, when aggregated *a posteriori* to coarse resolutions.

5.4 DISCUSSION

It is clear that any attempt to choose among the 4D-VAR and EnSRF DA approaches for solving a linear Gaussian problem, such as the CO_2 flux estimation problem, should be based on: (a) the tradeoff between slow convergence of the minimization algorithm (for 4D-VAR) or sampling error (for EnSRF) on the estimated fluxes, and (b) the carbon science questions being targeted. Given the availability of an adjoint model, and if only the best estimates are desired (for e.g., regional/continental budgets), the 4D-VAR is preferred over EnSRF, especially when the measurement network is highly variable. With a small number of iterations, the 4D-VAR may not converge to the global minimum but still reliably capture the majority of the large-scale features, with the final estimates close to those expected from a regular batch solution. The main disadvantages of the 4D-VAR is its incapability to provide any direct estimates of the analysis error and its more cumbersome implementation.

In the future, the 4D-VAR framework may be potentially more advantageous for the CO_2 flux estimation problem due to its ability to easily account for correlated observation errors (O. Talagrand, *pers. comm*.). Although this specific issue has not been demonstrated in this study, it is worth keeping in mind with the increasing use of satellite-based CO_2 measurements. Recent work by *Brankart et al*. [2009] has demonstrated techniques to cope with such correlations in an ensemble filter setting. Such techniques, however, are numerically efficient only for certain types of error correlation structures.

In the absence of correlated errors, the serial EnSRF is easier to implement and does not require the development and maintenance of an adjoint model. But due to restrictions on the size of the ensemble, it is necessary to adapt and tune additional algorithms like localization and inflation. Unlike the adaptive inflation piece, the localization component remains highly

subjective. Experiments in this study clearly showed that the localization length scale is dependent on both the ensemble size and the observational density. Will higher and higher volumes of observations push us towards specifying shorter localization length scales? If so, what is the limit beyond which decreasing the localization length scale may actually degrade the analysis? It is necessary to identify more rigorously a basis for selecting the localization parameters or develop approaches that may be less sensitive to changes in the observational network.

Ultimately the inflation and localization algorithms make the ensemble approximation to the full-rank Kalman Filter computationally feasible, and allow analysis of a fair quality. In spite of having more tuning parameters (ensemble size, localization length scale), the EnSRF formulation explicitly provides second-order statistical moments for the estimated system states. In an inverse problem framework, this is specifically desirable to ascertain the reliability of the estimates. The EnSRF is more desirable for attribution purposes, whereby source/sink estimates with confidence bounds can be used to gain a better understanding of the mechanistic processes driving the carbon cycle or reconcile estimates from top-down and bottom-up biospheric models. Recovering realistic *a posteriori* uncertainty bounds on the flux estimates will allow better tuning of the parameters in a bottom-up model, or at the minimum aid us in identifying a range of suitable values for these parameters.

With both approaches there is a direct trade-off between computational savings and estimation accuracy, which is intensified in trying to solve an under-determined problem with inconsistent observational network. For large-scale flux estimation problems, operational constraints will always exist along with scarce and inconsistent set of observations, erroneous transport models (thus further limiting the use of available observations) etc. The HT scheme

with limited number of ensemble members/descent iterations (Panel CO in Figures 5.3 and 5.4) serve as the closest analogue to how a real inversion problem is tackled. Even if we account for the increase in remote-sensing measurements of CO_2, the observational network is going to be a complex hybrid between the REF and the HT scheme. In this scenario, the accuracy and precision of either of the DA approaches are lower than the BIM for solving the CO_2 inverse problem until and unless the issues of convergence (4D-VAR) and sampling error (EnSRF) are appropriately addressed.

5.5 SUMMARY

This paper has compared two well-established data assimilation methods, 4D-VAR and EnSRF, in the context of a CO_2 flux estimation problem and investigated the potential impact of the estimation scale versus ensemble size (for the ensemble approach) and the number of descent iterations (for the variational approach). To assess the performance of the data assimilation approaches with the batch inverse modeling scheme, the inversion framework is designed in a relatively simple way. A 1-dimensional tracer transport problem is employed with the observational network and the overall inversion system prescribed analogous to a real CO_2 flux estimation problem. Both assimilation schemes were compared against the estimates from a synthesis BIM to isolate the degradation specifically due to the numerical approximations in the two approaches.

The sensitivity experiments demonstrated that when a large number of ensemble members or descent iterations are specified, state of the art implementation of the 4D-VAR and the EnSRF yield comparable analysis to the BIM, irrespective of the observational characteristics. When the operational parameters of the DA approaches are reduced, there is a complex interplay between the observations and the numerical approximations. Although both

methods are adept at handling the under-determined CO_2 flux estimation problem, their relative use should clearly be based on: a) the carbon science questions being targeted, and b) the inversion conditions under which they are being applied. By identifying the advantages and disadvantages of the two DA methods relative to the BIM, we expect that our finding will provide guidance to the CO_2 community in deciding which approach to invest in. Recently, each of these methods has begun to influence the development of the other, and it is expected that in the near future hybrid approaches may take center-stage for solving the CO_2 flux problem.

The conclusions about the operational implementation and the uncertainty estimates are well-established in the weather prediction community. To the best knowledge of the authors the sensitivity of the approaches to the observational schemes in the absence of an explicit dynamical model and for solving an under-determined inverse problem, has not been documented. The conclusions from our study can be juxtaposed against previous NWP-based inter-comparison studies, all of which have typically assumed a dynamical model. This study did not prescribe any form of a dynamical model and the findings are relevant to DA applications where dynamical models for propagating the state vector may not be readily available.

This study only considered simulated observations without any temporal correlation in the errors. The next step in the inter-comparison will be in the context of the assimilation of real remote-sensing observations. Controlled experiments with assimilation of satellite CO_2 data will be needed to fully understand the analyses quality in the future. How different DA algorithms account for these error sources and correlations within the remote-sensing data remain important research issues, not just within the CO_2-DA community but within the wider DA community as well.

Chapter 6

Towards Reliable Ensemble Kalman Filter Estimates of CO_2 Fluxes

6.1 INTRODUCTION

Over the last decade it has become increasingly apparent that quantification of global carbon sources and sinks with sufficient accuracy and precision is critical to balancing the global carbon budget and monitoring of carbon-management activities (*Schimel* [2007]). It has also become clear that our understanding of, and ability to accurately model, the carbon-cycle is severely constrained by the sparse distribution of the present atmospheric CO_2 measurement network (e.g. *Scholes et al*. [2009]). The sparse and spatially non-uniform network is neither sufficient to constrain regional budgets with the needed certainty, nor understand the nature, geographic distribution and temporal variability of CO_2 sources and sinks. This absence of spatially and temporally dense measurements of atmospheric CO_2 has spurred the development of space-based measurement sensors. Measurements from passive sensors such as the Atmospheric Infrared Sounder (AIRS) on Aqua (*Chahine et al*. [2006]), the Tropospheric Emissions Spectrometer (TES) on Aura (*Kulawik et al*. [2010]), the Infrared Atmospheric Sounding Interferometer (IASI) on Met-Op-1 (*Crevoisier et al*. [2009]), the SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY) on EnviSAT (*Buchwitz et al*. [2005]), the Greenhouse gases Observing SATellite (GOSAT; *Kuze et al*. [2009]), as well

as planned future sensors such as the Orbiting Carbon Observatory-2 (OCO-2; *Eldering et al*. [2012]), and the Active Sensing of CO_2 Emissions over Nights, Days, and Seasons (ASCENDS) satellite (*NRC* [2007]) are expected to improve our scientific understanding of regional carbon cycle processes and budgets. Although remote-sensing measurements of CO_2 do not achieve the precision possible from *in situ* measurements (*Rayner and O'Brien* [2001]), they provide a large number of observations with near-global coverage, beyond what is possible from a surface network alone (e.g. *Buchwitz et al*. [2007]).

The global coverage provided by these space-based measurements has demonstrated promise in improving the accuracy and precision of regionally-resolved flux estimates (e.g. *Baker et al*. [2010]), but the solution of the associated inverse problem has also resulted in a substantial increase in computational cost. The computational challenge results from the fact that inverse modeling techniques (a.k.a. top-down approaches) have historically been solved in "*batch*" mode, where the inversion is performed by solving a system of linear equations relating the CO_2 fluxes and the atmospheric CO_2 observations (e.g. *Enting* [2002]). Solving the batch problem requires an atmospheric transport model to be run either once per estimated flux region/period combination, or once per observation if an adjoint to the transport model is available. This becomes computationally infeasible given the increasing spatial and temporal resolution at which CO_2 fluxes are being estimated, and the increasing number of concentration measurements available from remote-sensing observations.

To address the increasing computational challenge of atmospheric inversions, data assimilation (DA) techniques (e.g. ensemble Kalman filter methods - *Peters et al*. [2005]; *Feng et al*. [2009]; *Miyazaki et al*. [2011]; variational methods - *Rayner et al*. [2005]; *Chevallier et al*. [2005b]; *Rödenbeck* [2005]; *Baker et al*. [2006a], or hybrid approaches such as the Maximum

Likelihood Ensemble Filter - *Zupanski et al*. [2007a]; *Lokupitiya et al*. [2008]) have recently been employed for estimating CO_2 fluxes, in some cases as part of advanced systems where meteorological and carbon variables are simultaneously assimilated (e.g. *Kang et al*. [2011]). Application of data assimilation techniques to the CO_2 problem is, however, much less mature (*Rayner* [2010]) than its use in numerical weather prediction (*Swinbank* [2010] and references therein) or for assimilating other atmospheric constituents such as humidity and ozone (e.g. *Rood* [2005]; *Lahoz and Errera* [2010]). An important question for carbon-cycle research that has hitherto remained unanswered concerns the impact of the numerical data assimilation framework on the precision and accuracy of fine-scale flux estimates and their associated uncertainties. Previous CO_2 DA studies have evaluated flux estimates by comparing them to biospheric model and inventory estimates and/or by assessing how well they reproduce available atmospheric CO_2 observations (e.g. *Peters et al*. [2005]; *Chevallier et al*. [2007]; *Lokupitiya et al*. [2008]; *Feng et al*. [2009]; *Baker et al*. [2010]; *Kang et al*. [2011]; *Miyazaki et al*. [2011]). Given the host of error sources (e.g. transport error, aggregation error, etc.) that impact inversions, these diagnostics provide an assessment of the overall inversion framework, but do not isolate any errors incurred due to the numerical approximations in the implemented DA approach.

This study is primarily motivated by an attempt to isolate and quantify such errors, specifically from the perspective of an ensemble Kalman filter applied to the estimation of CO_2 fluxes at fine spatial and temporal scales. Ensemble filters and their variants (e.g. *Peters et al*. [2005, 2007, 2010]; *Zupanski et al*. [2007]; *Lokupitiya et al*. [2008]; *Feng et al*. [2009, 2011]; *Kang et al*. [2011]; *Miyazaki et al*. [2011]) have gained popularity within the carbon-cycle community due to their simple conceptual formulation and relative ease of implementation. So far, the examination of the use of ensemble Kalman filters for estimating fluxes at fine spatial

and temporal scales has been limited, however. Except for *Kang et al*. [2011] and *Miyazaki et al*. [2011], where fluxes were estimated at the grid resolution of the atmospheric transport model used in the studies (~2.8°), almost all other studies have estimated fluxes at large spatial scales (e.g. continental or ecoregion). The temporal scales at which fluxes have been estimated range from several days to weeks.

The work presented here estimates fluxes at substantially finer scales (1° by 1° and 3-hourly) relative to previous application of ensemble filters. In general, high resolution estimates of carbon fluxes are advantageous for 1) improving budgeting and mechanistic understanding of the carbon cycle at local to regional scales, and 2) minimizing spatial and temporal aggregation errors (e.g. *Kaminski et al*. [2001]; *Peters et al*. [2010]; *Gourdji et al*. [2010]) that may otherwise bias the final flux estimates. The impact of spatial aggregation errors has long been discussed and documented in the inverse modeling literature (e.g. *Kaminski et al*. [2001]; *Engelen et al*. [2002]; *Peters et al*. [2010]), and recent studies have shown that *a priori* temporal aggregation has similar impacts. *Gourdji et al*. [2010; 2012] found that biases occurred when fluxes were estimated at multi-day or even daily timescales, and recommended a 3-hourly temporal resolution to allow the inversion to resolve the diurnal cycle. *Huntzinger et al*. [2011] further found that differences between the diurnal representations among a suite of terrestrial ecosystem models yielded significant difference at CO_2 monitoring locations, suggesting that adopting a fixed diurnal cycle from one particular model *a priori* could bias flux estimates at larger scales. Although desirable from a scientific perspective, applying an ensemble approach to a fine-scale flux estimation problem is challenging due to two issues associated with the ensemble filter.

The first challenge common to all applications of ensemble filters is the error due to representing the probability density function of the fluxes by a finite number of randomly

generated flux realizations or system states. Experience in the NWP area has suggested that because of the finite number of ensemble members, the ensemble filter can suffer from variance underestimation, rank deficiency and sampling error (e.g. *Houtekamer and Mitchell* [2005]; *Anderson* [2007a; 2007b]; *Ehrendorfer* [2007]; *Meng and Zhang* [2011]), all of which impact both the final estimates and their uncertainty. *Anderson* [2007a] notes that, even in low-order perfect-model applications for NWP, mitigating the impacts of the limited ensemble size requires the introduction and tuning of several additional algorithms. Considerable expertise exists in these algorithms for NWP (e.g. *Hamill and Whitaker* [2005]; *Anderson* [2007a; 2007b]; *Uzunoglu et al*. [2007]; *Sacher and Bartello* [2008]; *Anderson* [2009]; *Bergemann and Reich* [2010]; *Bishop and Hodyss* [2011]) and for DA of other constituents (e.g. *Schutgens et al*. [2010]). However, these algorithms and their impact on flux estimates and uncertainties are less well-understood for the carbon flux estimation problem. Applications to CO_2 have investigated the impact of ensemble size (e.g. *Peters et al*. [2005]; *Zupanski et al*. [2007a]; *Feng et al*. [2009]; *Miyazaki et al*. [2011]) and different localization/inflation parameters (see Sections 6.2.3 and 6.2.4) but have refrained from drawing conclusions as to the optimal values of parameters that may aid future filter designs. Because previous studies have not compared ensemble filter estimates to those from batch inversions, it is difficult to isolate the impact of the parameter and algorithm choices from other errors present within any inversion framework.

The second challenge that differentiates the CO_2 flux estimation problem from NWP-related applications is that there is currently no dynamical model to directly evolve the carbon flux state vector forward in time (*Peters et al*. [2005]; *Lokupitiya et al*. [2008]; *Miyazaki et al*. [2011]). In other words, given the estimated flux at one time, there is no model to predict the flux at the following assimilation time. The lack of such a dynamical model represents a loss of

valuable information to the ensemble, as along with the transport model, a dynamical model would capture the flow-dependent error covariance patterns. In NWP-related applications, several studies have been carried out to test the impact of dynamical model errors (e.g. *Houtekamer et al*. [2005]; *Szunyogh et al*. [2005]; *Houtekamer et al*. [2009]; *Hamill and Whitaker* [2011]), but no study has evaluated the impact of a complete absence of a dynamical model. The absence of the dynamical model may make the ensemble filter extremely sensitive to the observation network and coverage. Given the spatial and temporal variability of atmospheric CO_2 measurements (whether in situ or satellite-based), this raises questions about the applicability of ensemble filters for leveraging the information content of available CO_2 observations.

In order to understand these issues, we introduce a geostatistical variant of the Bayesian ensemble square root filter (EnSRF; *Whitaker and Hamill* [2002]). The geostatistical ensemble square root filter (GEnSRF) is based on a geostatistical inverse modeling formulation of the flux estimation problem (GIM; *Michalak et al*. [2004]). The GIM formulation is not limited to the use of prior CO_2 flux information from biospheric models and/or inventories, and has been applied for inversions conducted at very high spatiotemporal resolutions (e.g. *Gourdji et al*. [2012]).

The GEnSRF is used as a prototype filter in exploring the impacts of the challenges outlined above for CO_2 flux estimates at fine spatial and temporal scales. The sensitivity of the ensemble filter to different scenarios is judged by comparing the GEnSRF estimates to the estimates from an equivalent batch GIM setup. This comparison is carried out using synthetic data from the growing season (June 2008) over North America. Both GEnSRF and GIM are used to estimate fluxes and their associated uncertainties at a $1^o \times 1^o$ (spatial) and 3-hourly (temporal) resolution. Test cases are designed to gauge the performance of the ensemble system and to

evaluate whether the numerically-approximate ensemble scheme can accurately capture the characteristic features of the CO_2 cycle, such as the spatial location of sources and sinks and the amplitude and phase of the diurnal flux cycle. The test cases are used to assess the baseline performance of the ensemble system, as well as to explore the impact of the measurement network, ensemble size, and the implementation of covariance inflation and localization algorithms designed to improve ensemble performance.

Overall, this work provides 1) an assessment of the relative performance of the ensemble filter in comparison to the batch approach and of the conditions necessary for the ensemble approach to be a suitable replacement for batch inversions, and 2) an investigation of the error sources in the ensemble system and their implications for adjustments to ensemble systems that need to be made relative to NWP applications. The remainder of this paper is organized as follows. Section 6.2 provides the rationale for the proposed filter followed by an overview of GEnSRF. Section 6.3 provides a description of the examined synthetic data test cases. Results are presented and discussed in Section 6.4. Finally, we conclude in Section 6.5 with a summary of the findings of this study and recommendations for future research.

6.2 METHODOLOGY

6.2.1 Choosing a filter formulation

The underlying framework in all ensemble filters is a low-rank ensemble representation of the error covariance matrices. The ensembles themselves are scaled matrix square-roots of the covariance matrices, and are updated during the assimilation of observations either stochastically (e.g. *Houtekamer and Mitchell* [1998]; *Burgers et al*. [1998]; *Pham* [2001]) or deterministically (e.g. *Bishop et al*. [2001]; *Anderson* [2001]; *Whitaker and Hamill* [2002]; *Ott et al*. [2004]). The

details of this update step distinguish most ensemble variants. Based on existing studies (e.g. *Tippett et al*. [2003]; *Lawson and Hansen* [2004]; *Nerger et al*. [2005]; *O'Kane and Frederiksen* [2008]), it can be concluded that for a linear problem – 1) deterministic filters are more accurate than their stochastic counterparts, and 2) although all the deterministic filters will produce analysis ensembles that span the same state subspace and have the same covariance, implementation of a serial EnSRF has the lowest computational cost if the observation errors are assumed to be independent.

The simplest serial EnSRF that can be implemented for inferring CO_2 surface fluxes is one using a Bayesian formulation (e.g. Carbon Tracker - *Peters et al*. [2010]), which uses prior information about the CO_2 fluxes from bottom-up models and/or inventories. Because of the highly ill-posed nature of the CO_2 flux estimation problem and the sparseness of the current observing network, the posterior flux estimates and uncertainties are quite sensitive to the *a priori* prescribed flux patterns and their associated error covariance parameters (*Peters et al*. [2010]). By adapting the ensemble system to the geostatistical approach, we avoid some of the reliance on prior/model assumptions, albeit at the cost of an increase in complexity. Therefore, it can be argued that the niche filled by a geostatistical ensemble square root filter lies in more directly isolating the information content of the available atmospheric measurements.

6.2.2 Geostatistical ensemble square root filter

The GEnSRF, like the EnSRF, is a Monte Carlo technique based on a state space formulation of the Kalman Filter using an ensemble of model states to represent, propagate and update the estimates of the state and state error covariance. The aim is to minimize an objective function of the form:

$$J(\mathbf{s},\boldsymbol{\beta}) = \frac{1}{2}\left[\mathbf{z} - h(\mathbf{s})\right]^T \mathbf{R}^{-1}\left[\mathbf{z} - h(\mathbf{s})\right] + \frac{1}{2}\left[\mathbf{s} - \mathbf{X}\boldsymbol{\beta}\right]^T \left(\mathbf{Q}^b\right)^{-1}\left[\mathbf{s} - \mathbf{X}\boldsymbol{\beta}\right] \quad 6.1$$

where **z** is a *n*×1 vector of observations, *h* represents the atmospheric transport model, **s** is a *m*×1 state vector composed of the discretized *unknown* surface flux distribution, **R** is the *n*×*n* model-data mismatch covariance, **X** is an *m*×*1* vectors of ones in the test cases presented in the work, but could also include auxiliary variables related to carbon flux (see *Gourdji et al*. [2008; 2012] for further details on the selection of auxiliary variables), **β** is an unknown constant here, but could also include unknown drift coefficients that scale the auxiliary variables in **X**, the prior covariance matrix $\mathbf{Q}^b$ describes the expected variability in flux departures from **Xβ** as a function of the separation distance in space and time between fluxes (see Section 6.3.1.3 for further details on the structure of $\mathbf{Q}^b$), and *T* represents the transpose operator. In a batch setup, instead of running the transport model *h* directly as part of the inversion, an *n*×*m* sensitivity matrix **H** (a.k.a. Jacobian matrix) is generated that represents the sensitivity of the observations **z** to the fluxes **s** (i.e., $H_{i,j} = \partial z_i / \partial s_j$).

Equation 6.1 represents a compromise between reproducing the atmospheric measurements (**z**) and staying close to the statistical model of the trend (**Xβ**), where the covariance matrices determine the relative weight of these competing objectives. Although some implementations of ensemble approaches include more variable in the state vector **s**, including atmospheric concentrations of CO_2 themselves, the focus here is on constraining only the underlying fluxes. Correspondingly, any updates in the atmospheric CO_2 distribution must therefore be attributable by a change in the underlying fluxes.

The GEnSRF is implemented as a smoother (e.g. *Bruhwiler et al*. [2005]; *Michalak* [2008]), such that individual time steps through the smoother include a) fluxes that are no longer being estimated, b) fluxes that are being updated at least for the second time (i.e., that have been

previously estimated), and c) fluxes being estimated for the first time (i.e., for which no prior information is available). In the following discussion, the *m*×1 vector of the *estimated* surface flux distribution is represented as $\hat{\mathbf{s}}$. $\hat{\mathbf{s}}_j$ denotes estimates of fluxes that are being updated at least for the second time, $\hat{\mathbf{s}}_k$ denotes fluxes that are being obtained for the first time, and $\hat{\mathbf{s}}_{j+k}$ denotes both sets of fluxes being estimated. Finally, the superscripts *a* and *b* represents the analyzed (or updated) estimate and the previous estimate.

Given an initial prior covariance $\mathbf{Q}^b$, GEnSRF starts by creating an ensemble of *N* state fields (where *N*<<*m*). These are created as unconditional realizations of the matrix $\mathbf{Q}^b$ through Cholesky decomposition.

$$\mathbf{Q}^b \approx \frac{1}{N-1}\left(\mathbf{s}'^b\right)\left(\mathbf{s}'^b\right)^T \tag{6.2}$$

where $\mathbf{s}'^b$ represents the estimated error statistics of CO_2 flux deviations from the trend. In the limit of $N \to \infty$ this representation of $\mathbf{Q}^b$ is exact. In GEnSRF, observations are assimilated serially. When the i^{th} observation is being assimilated, the estimates of fluxes are given by:

$$\begin{bmatrix} \hat{\mathbf{s}}_j^a \\ \hat{\mathbf{s}}_k^a \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{s}}_j^b \\ 0 \end{bmatrix} + \mathbf{\Lambda}\left[\mathrm{z}_i - h_i\left(\hat{\mathbf{s}}_j^b\right)\right] \tag{6.3}$$

where $\mathbf{\Lambda}$ is calculated by solving the system of equations,

$$\begin{bmatrix} \mathbf{H}_{i,j+k}\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k} + \mathrm{R}_{i,i} & h_i\left(\mathbf{X}_k\right) \\ \left(h_i\left(\mathbf{X}_k\right)\right)^T & 0 \end{bmatrix}\begin{bmatrix} \mathbf{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \left(\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k}\right)^T \\ \begin{bmatrix} \mathbf{0} & \mathbf{X}_k^T \end{bmatrix} \end{bmatrix} \tag{6.4}$$

Consistent with a GIM framework, fluxes being estimated for the first time $\hat{\mathbf{s}}_k$ need not be initialized with a prior value, and the zero in equation 6.3 is not equivalent to a prior in the classical Bayesian setup. Since direct matrix computation of $\mathbf{H}_{i,j+k}\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k}$ and

$\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k}$ can be expensive, these are approximated by running the transport model directly with the ensemble of state deviations.

$$\mathbf{H}_{i,j+k}\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k} \approx \frac{1}{N-1}\left(h_i\left(\mathbf{s}'^b_{j+k}\right)\right)\left(h_i\left(\mathbf{s}'^b_{j+k}\right)\right)^T \tag{6.5}$$

$$\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k} \approx \frac{1}{N-1}\left(\mathbf{s}'^b_{j+k}\right)\left(h_i\left(\mathbf{s}'^b_{j+k}\right)\right)^T \tag{6.6}$$

Once $\mathbf{\Lambda}$ is obtained, it is used in Equation 6.3 to estimate the fluxes. If the same $\mathbf{\Lambda}$ is used to update the ensemble of state deviations from the mean $\mathbf{s}'^b_{j+k}$, it would result in an underestimation of the analysis error covariance (*Whitaker and Hamill* [2002]). Instead $\mathbf{\Lambda}$ is reduced in magnitude ($\tilde{\mathbf{\Lambda}}$; Equation 6.7) such that the spread of the ensemble is reduced less by the analysis (Equation 6.8), in order to maintain an error covariance consistent with the full-rank Kalman filter.

$$\tilde{\mathbf{\Lambda}} = \left(1+\sqrt{\frac{\mathrm{R}_{i,i}}{\mathbf{H}_{i,j+k}\mathbf{Q}^b_{j+k,j+k}\mathbf{H}^T_{i,j+k}+\mathrm{R}_{i,i}}}\right)^{-1}\mathbf{\Lambda} \tag{6.7}$$

$$\mathbf{s}'^a_{j+k} = \mathbf{s}'^b_{j+k} - \tilde{\mathbf{\Lambda}}\left(h_i\left(\mathbf{s}'^b_{j+k}\right)\right) \tag{6.8}$$

When observations are serially processed, Equation 6.7 reduces to the computation of a scalar factor. Notice that the piece $h_i\left(\mathbf{s}'^b_{j+k}\right)$ is already available, and hence updating the ensemble via Equation 6.8 is no more computationally expensive than Equation 6.3.

Finally, before assimilating the next observation, we update the sampled observational ensemble and the sampled observation state corresponding to all future observations that are yet to be assimilated:

$$h_{i+1}\left(\mathbf{s}'^{a}\right) = h_{i+1}\left(\mathbf{s}'^{b}\right) - h_{i+1}\left(\tilde{\mathbf{\Lambda}}\, h_i\left(\mathbf{s}'^{b}_{j+k}\right)\right) \qquad 6.9$$

$$h_{i+1}\left(\hat{\mathbf{s}}^{a}\right) = h_{i+1}\left(\hat{\mathbf{s}}^{b}\right) + h_{i+1}\left(\mathbf{\Lambda}\left[\mathbf{z}_i - h_i\left(\hat{\mathbf{s}}^{b}_{j}\right)\right]\right) \qquad 6.10$$

Equations 6.9 and 6.10 require two additional transport model runs that could be avoided as described in *Peters et al*. [2005] by approximating these equations in a manner similar to Equation 6.5. When the ensemble size is much smaller than the size of the state vector (as will be the case when fluxes are estimated at a high spatial and temporal resolution) this results in a poor approximation, however. The additional cost of running the transport model might well offset the errors incurred due to the approximation. However, with satellite measurements, running the transport model each time an observation is assimilated makes the direct implementation of Equations 6.9 and 6.10 impractical. Hence, work is underway to find suitable alternatives to both these equations without incurring large errors in the analysis. In the work presented here, these additional runs are performed.

Using Equations 6.3 to 6.10, a best estimate of CO_2 fluxes is obtained. Once all observations have been assimilated, the *a posteriori* covariance $\mathbf{Q}^a$ for the flux estimates is reconstructed from the ensemble (Equation 6.11). The diagonal values of this posterior covariance matrix correspond to the uncertainty (expressed as a variance) of each estimated flux component in $\hat{\mathbf{s}}$.

$$\mathbf{Q}^a = \frac{1}{N-1}\left(\mathbf{s}'^{a}\right)\left(\mathbf{s}'^{a}\right)^T \qquad 6.11$$

Finally, in regular NWP applications a dynamical model (non-linear forecast operator) would have been used to propagate the state vector between the two observational time periods. For the CO_2 problem, no suitable deterministic model is available to directly propagate fluxes

from one time step to the next. This differentiates the CO_2 flux estimation problem from the NWP and other trace gas assimilation problems, and may have critical implications for good filter performance. Note that this drawback is not specific to GEnSRF but to all variants of the ensemble filter that have been employed for CO_2 flux estimation.

6.2.3 Covariance localization

Covariance localization aims to heuristically improve the error covariance estimates in the case of small ensemble sizes. In all ensemble filters (including GEnSRF), the number N of ensemble members is small relative to the size m of the state space, hence the representation of the prior covariance matrix in N-dimensional space is not perfect. This results in a number of erroneous flux correlations as a consequence of which a state variable may be incorrectly impacted by an observation that is physically remote.

Several covariance localization techniques have been proposed for the NWP problem (*Houtekamer and Mitchell* [2001]; *Hamill et al.* [2001]; *Ott et al*. [2004]; *Anderson* [2007b]) to account for the statistical noise of the ensemble. For CO_2 applications, implemented localization schemes have varied depending on the particular ensemble filter variant being used. For example, *Peters et al*. [2005] chose a simple exponential decay function, while *Miyazaki et al.* [2011] subjectively specify different cutoff radii based on the type and location of observation data used in their analysis. *Zupanski et al*. [2007a] and *Lokupitiya et al*. [2008] chose a more dynamic scheme based on information theory, where the localization length scale is a function of the information content in the assimilated observations.

Similarly to *Peters et al*. [2005], we implement a simple covariance localization scheme in GEnSRF. This is achieved by performing a Schur (or Hadamard; *Horn and Mathias* [1990])

product, or element-wise multiplication (denoted • in Equation 6.12) of a correlation matrix ρ with the covariance model generated by the ensemble as shown in Equation 6.12.

$$\mathbf{Q}^{b}_{j+k,\,j+k}\mathbf{H}^{T}_{i,\,j+k} \approx \rho \bullet \left[\frac{1}{N-1}\left(\mathbf{s}'^{b}_{j+k}\right)\left(h_i\left(\mathbf{s}'^{b}_{j+k}\right)\right)^{T} \right] \qquad 6.12$$

Here, ρ is defined using a standard fifth-order Gaspari-Cohn function (*Gaspari and Cohn* [1999]) with a finite length scale. Note that *Peters et al*. [2005] used an exponential decay function to define their ρ. Both the Gaspari-Cohn function and the exponential function are compactly supported (*Gneiting* [2002]; *Bergemann and Reich* [2010]), which means that the function is non-zero in only a small (local) region specified by a length scale. We find that the overall conclusions presented in Section 6.4.2 are valid for a variety of compactly supported functions. The key ingredient in all compactly supported functions is the length scale, which ensures that spurious correlations are removed, but correctly specified physical correlations are maintained and not excessively damped.

Covariance localization using the Schur product might be a simple approach to increase the effective rank of the covariance matrix, but there are several important caveats for CO_2 applications. Previous studies (e.g. *Lokupitiya et al*. [2008]) have raised questions regarding selection of an appropriate localization length scale, and whether the atmospheric advection of CO_2 is consistent with the use of a compactly supported correlation function such as the Gaspari-Cohn. Since the prior covariance matrix holds information on the spatial and temporal autocorrelation of flux deviations from the trend (Section 6.3.1.3), by including the Schur product and thereby modifying this matrix, covariance localization may disrupt the autocorrelation structure (see *Karspeck and Anderson* [2007]; *Oke et al*. [2007]; *Kepert* [2009] for a similar discussion related to NWP problems). In spite of these concerns, in this study, we have persisted with the Gaspari-Cohn function because we want to assess the applicability of this

simple scheme for atmospheric CO_2 inversions. Future work could explore the applicability of more sophisticated dynamic localization schemes (e.g. *Zupanski et al*. [2007a]) or balance-aware localization schemes proposed for NWP (e.g. *Bishop and Hodyss* [2011]; *Kepert* [2011]; *Jun et al*. [2011]) or more adaptive techniques based on the prior ensemble (e.g. *Anderson* [2012]).

6.2.4 Adaptive covariance inflation

A second algorithm for combatting insufficient variance in ensemble filters is covariance inflation. Insufficient variance (or under sampling) is primarily caused by sampling error resulting from the use of small ensembles (*Furrer and Bengtsson* [2007]). Over successive assimilation periods, under- sampling can become more severe, and in the worst case can lead to filter divergence, where the filter effectively rejects the observations and the assimilation reduces to the prior. Also, because the posterior analysis error covariance is generated from the ensemble at the end of the assimilation (Equation 6.11), insufficient variance leads to an under-estimation of the analysis error covariance (i.e. the flux uncertainties).

Several ad-hoc and adaptive techniques have been proposed in NWP applications to counter this loss of variance (e.g. *Anderson and Anderson* [1999]; *Zhang et al*. [2004]; *Hamill and Whitaker* [2005]; *Anderson* [2007a]; *Li et al*. [2009]; *Anderson* [2009]; *Pena et al*. [2010]). Various inflation schemes have also been employed with CO_2 applications, depending on the variant of ensemble Kalman filter used in a particular study. Some have avoided using the ensemble spread as a measure of uncertainty altogether by instead deriving final uncertainties from a set of sensitivity experiments (e.g. *Peters et al*. [2010]). *Feng et al*. [2009] chose to use an ensemble of the same size as the state vector, thereby minimizing under sampling directly. For the maximum likelihood ensemble filter (e.g. *Zupanski et al*. [2007a]; *Lokupitiya et al*. [2008]) a multiplicative inflation scheme was used for covariance inflation, where the ensemble

is inflated by a constant factor that is homogeneous in space (although different inflation factors are used for land and ocean regions) and time. This approach has some limitations because neither the observation network nor the CO_2 dynamics are homogeneous in space and time, and the cost of tuning experiments to find an appropriate inflation factor that is applicable everywhere is prohibitive. Recognizing these drawbacks, more recent studies have employed either conditional covariance inflation (*Miyazaki et al*. [2011]) or a mix of adaptive and covariance relaxation techniques (*Kang et al*. [2011]).

In GEnSRF, we adopt the more generalized version of the adaptive technique used by *Kang et al*. [2011] (as originally proposed by *Anderson* [2009]) to calculate spatially and temporally varying inflation factors for each state component (i.e., flux at each time and grid point). This adaptive algorithm applies Bayesian estimation theory to the probability density function of the inflation factors. First, a normally-distributed inflation random variable is associated with each element of the state vector. Then, via Bayes theorem, these inflation factors are incrementally updated during serial assimilation of the observations. Note that the atmospheric CO_2 observations can be used to optimize the inflation factors for the CO_2 fluxes due to the link between these quantities provided by the atmospheric transport model.

In order to calculate the spatially and temporally varying inflation factors, however, it is necessary to implement covariance localization first. The adaptive technique uses sample correlations of the ensemble between observation space and the model space to convert the inflation estimates in the observation space to those in the model space. Covariance localization plays an important role in reducing the sampling noise in the sample correlations. If no covariance localization is pursued, then the sampling error manifests itself in the adaptive inflation step resulting in spurious inflation factors. Thus, using the adaptive technique, we have

specifically adjusted the covariance inflation strategy to take into account the information provided by the atmospheric CO_2 measurements.

We refer the reader to *Anderson* [2009] (and the subsequent review by *Miyoshi* [2011]) for the mathematical underpinnings of the adaptive approach. It is worthwhile to reiterate that this particular adaptive technique has not previously been applied to any CO_2 inversion study. Hence, as part of the sensitivity tests described later, we will examine both the advantages and disadvantages of this adaptive technique for the CO_2 source-sink estimation problem.

6.3 SAMPLE APPLICATION

The GEnSRF approach is applied to a synthetic data study over the North American continent (Section 6.3.1). A series of analyses are designed (Section 6.3.2) to compare the estimates from GEnSRF with the estimates from GIM. These comparisons are done by aggregating the posterior estimates to a range of spatial and temporal scales (Section 6.3.3) to assess the accuracy and precision of the ensemble approach relative to a batch inversion.

In the context of this study, the main advantage of the GIM approach relative to the GEnSRF technique is that it solves the entire system of equations analytically (without any approximations) and hence provides a "*gold*" standard for evaluating the ensemble results. By keeping the atmospheric datasets consistent for GEnSRF and GIM, it is possible to isolate the degradation due to the numerical approximations in the ensemble filter framework. The drawback of this setup is that the effects of the transport model errors have been removed by using the same transport model to both create the synthetic measurements as well as estimate the fluxes in the inversion.

6.3.1 Experimental design

6.3.1.1 Flux data and Basis functions

Biospheric fluxes from the Carnegie Ames Stanford Approach terrestrial carbon cycle model, as configured for the Global Fire Emissions Database v2 project (henceforth referred to as CASA-GFEDv2; *Randerson et al*. [1997]; *van der Werf et al*. [2006]) are used as the true fluxes for generating the synthetic atmospheric data. The monthly-averaged 1° × 1° CASA-GFEDv2 Net Ecosystem Exchange (NEE) for June 2008 (Figure 6.1) is temporally downscaled to 3-hourly resolution using the method of *Olsen and Randerson* [2004].

The sensitivity matrix **H** is obtained by coupling the Weather Research Forecasting (WRF) model (*Skamarock et al*. [2005]) to the Stochastic Time-Inverted Lagrangian Transport Model (STILT) (*Lin et al*. [2003]), as outlined in *Gourdji et al*. [2010]. Calculating and pre-storing the sensitivity matrix **H** is necessary for performing the batch GIM analysis but not for GEnSRF, where the transport model can be run directly as part of the DA system. Given that **H** was available in this case, it is also used as the transport model for the ensemble implementation.

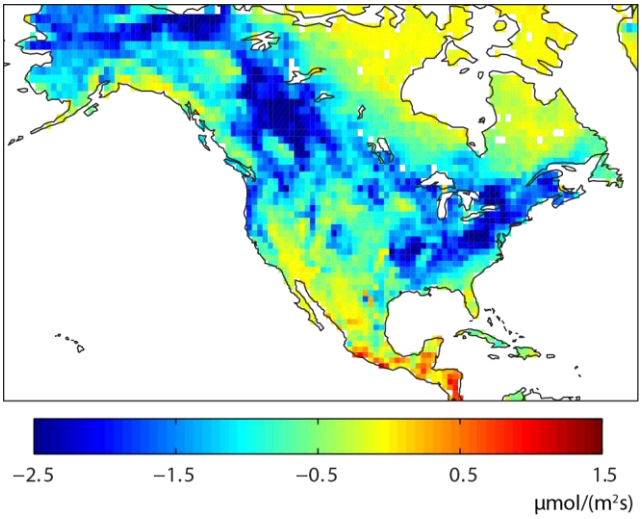


Figure 6.1 *- "True" CASA-GFEDv2 fluxes aggregated to the monthly scale.*

6.3.1.2 Synthetic observation data

The basis functions generated via WRF-STILT are used with the CASA-GFEDv2 fluxes to generate the synthetic observations **z** (i.e., $h(\mathbf{s})$) for the 35 continuous observation towers (see Appendix C) that were operational in June 2008 (Figure 6.2). First, a full set of synthetic data is generated for all the towers at the 3-hourly scale, and small random errors (standard deviation of 0.1ppm) are then added to the synthetic data. Such small errors were used to represent an, albeit somewhat unrealistic, best case scenario for the performance of the ensemble approach. Next only afternoon measurements are retained for the shorter towers (height ≤ 150 m) consistent with typical data choices in inversion studies (e.g. *Göckede et al*. [2010]; *Gourdji et al*. [2010]), motivated by lower transport model errors for afternoon conditions (e.g. *Geels et al*. [2007]; *Gerbig et al*. [2008]). Finally, data gaps are simulated in the synthetic observations consistent

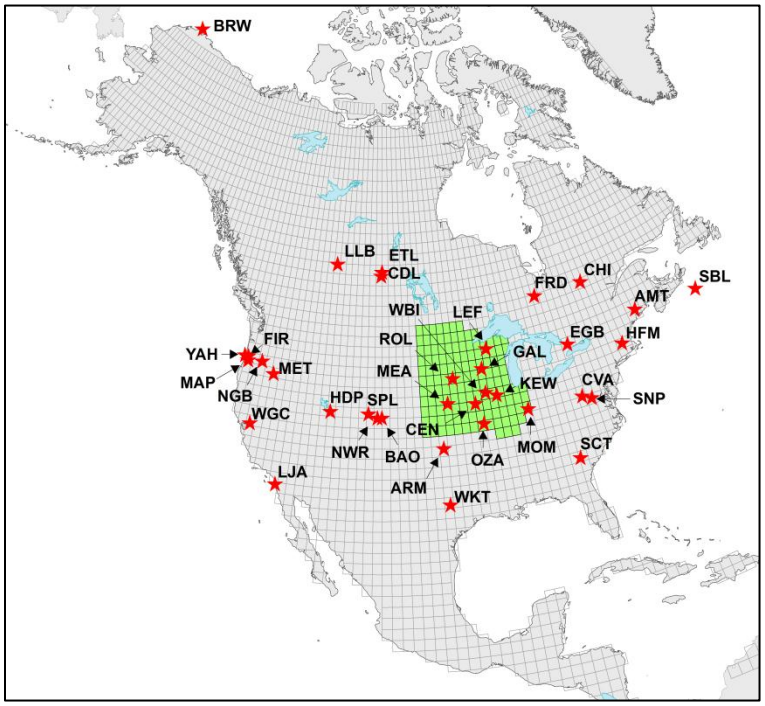


Figure 6.2 *- Location of the 35 tower network (stars), and the regions used for interpreting the flux estimates, i.e., North America and the MCI region (green shaded area). The background grid represents the flux estimation resolution of 1°× 1°. The three-letter codes for the towers are defined in Appendix C.*

with missing data from the actual June 2008 observations (due to either instrument down time or calibration needs). By mimicking the random outage in the collection/storage of the data, the synthetic dataset is highly variable (in both space and time), but realistic. The ratio of the number of fluxes (m) to the number of observations (n) is on the order of ~250:1. Conversely, if the full set (i.e., eight 3-hourly-averaged observations per day) of synthetic observations without data gaps were retained, then the ratio would be on the order of ~75:1.

6.3.1.3 Error covariance matrices

The model-data mismatch covariance matrix **R** is a diagonal matrix, with values of 0.01 ppm^2 along the diagonal (i.e., all towers are assumed to have the same model-data mismatch error), corresponding to the variance of the errors introduced into the synthetic observations.

The prior covariance matrix $\mathbf{Q}^b$ captures the spatiotemporal autocorrelation of the flux deviations from the model of the trend **Xβ.** In this study, only spatial correlation is assumed *a priori* in order to keep the structure of $\mathbf{Q}^b$ simple, although accounting for both spatial and temporal correlation could further improve estimates (e.g. *Gourdji et al*. [2010]; *Chevallier et al*. [2012]).

$\mathbf{Q}^b$ is prescribed as a block diagonal matrix, with each block describing the correlation between grid-scale fluxes for each time period of the inversion. Based on previous work (e.g. *Michalak et al*. [2004]; *Gourdji et al*. [2010]), each block is modeled by an exponential covariance function:

$$\mathbf{Q}^b\left(d \mid \sigma^2, l\right) = \sigma^2 \exp\left(-\frac{d}{l}\right) \qquad 6.13$$

where d is the spatial separation distance between the grid points where fluxes are to be estimated, σ^2 represents the variance of the flux residuals at large separation distances, and l is

the range parameter. The correlation length beyond which correlation between the flux residuals becomes negligible is approximately 3*l* (*Chiles and Delfiner* [1999]) for an exponential model.

The covariance parameters in $\mathbf{Q}^b$ (i.e., σ^2 and *l*) can be obtained via different methods (e.g. *Michalak et al*. [2005]; *Rödenbeck* [2005]; *Chevallier et al*. [2010]) ranging from analyzing the variability in biospheric model outputs to statistically inferring these parameters directly from the atmospheric measurements. In this study, we follow the latter approach, and optimize for the covariance parameters using the Restricted Maximum Likelihood (e.g. *Kitanidis* [1995]; *Michalak et al*. [2004]).

6.3.2 Test cases

Two primary inversion setups (TC1 and TC2) are outlined, both of which estimate 3-hourly fluxes at a 1° × 1° scale over North America for the month of June 2008. However, TC1 uses a sparse measurement dataset (as described in Section 6.3.1.2), while TC2 uses all 24 hours of measurements for all towers, yielding a temporally denser and homogeneous dataset. Covariance parameters were estimated separately for the two test cases. GIM is run once for each test case to obtain the batch estimates, while the GEnSRF is run multiple times for both TC1 and TC2 with different configurations of ensemble size, localization and inflation parameters. The details of these runs are expanded upon in the following paragraphs.

A control run of GEnSRF is defined based on TC1 with a 500-member ensemble and without any covariance localization or adaptive inflation. This run, designated as TC1E500, is used to gauge the incremental benefits of subsequent modifications.

Given the absence of a dynamical model to propagate the state vector forward in time, our hypothesis is that the inversion conditions (at least in terms of measurement quantity and density) may play a significant role in the performance of the ensemble filter. *Miyazaki et al*.

[2011] concluded that the absence of a dynamical model resulted in the posterior analysis being sensitive to the initial error covariance, but this earlier study did not test the influence of the measurement network. TC2 is designed to explore the impact of the measurement network sparseness as it represents the best possible scenario that one can attain with the existing ground-based continuous measurement network. If this were a real-data study, there would be several caveats regarding using all 3-hourly measurements, especially from shorter towers (*Gourdji et al*. [2012]). Hence other inversion scenarios in which measurements are progressively reduced in space and time were also evaluated. The conclusions from these additional test cases mirrored those from TC1 and TC2, and hence these have been omitted here for the sake of brevity.

In order to provide insight into an optimal and practical setup of the ensemble filter that can provide accurate flux estimates of CO_2 with reliable uncertainties, the parameters of the ensemble system were varied for both TC1 and TC2. The GEnSRF is run with three ensemble sizes – 100, 500, and 2500, denoted as E100, E500, and E2500, respectively. In addition, three different localization length scales were prescribed – 500, 1500 and 3000 km, denoted as L500, L1500 and L3000, respectively. Finally, the adaptive inflation algorithm requires *a priori* estimates of inflation factors and their associated variance. Again three different specifications of the prior inflation variance were provided - a prior inflation factor of 1 with a standard deviation of 0.01, a prior inflation factor of 1 with a standard deviation of 0.05, and a prior inflation factor of 1 with a standard deviation of 0.25. These runs are denoted as I001, I005 and I025, respectively. These parameters were chosen based on a combination of extensive literature review of ensemble filter applications, subjective knowledge of CO_2 transport and its correlation scales, and some preliminary testing with a 1D advection-diffusion problem.

Overall, a total of 2 GIM and 56 GEnSRF runs were carried out for this study. The 2 GIM runs represent the batch estimates for each of the test cases, and are simply denoted as GIM TC1 and GIM TC2. For GEnSRF, the first run for each setup is with an ensemble size of 500 and without any localization and inflation applied (i.e., the control run GEnSRF TC1E500 and GEnSRF TC2E500). GEnSRF is then run with varying ensemble sizes, localization length scales, and different *a priori* inflation values as described above. As an example, the GEnSRF run for setup TC1 using a 500 member ensemble, a localization length scale of 1500 km and a prior inflation factor of 1 with a standard deviation of 0.05 is denoted GEnSRF TC1E500_L1500I005.

Finally, as mentioned previously, the covariance parameters in $\mathbf{Q}^b$ (Equation 6.13) are optimized separately for the two inversion setups. The flux standard deviation (σ) was 6.7 $\mu mol*m^{-2}s^{-1}$ for TC1 and 6.4 $\mu mol*m^{-2}s^{-1}$ for TC2, while the correlation length ($3l$) was 1630 km for TC1 and 1590 km for TC2. For all the runs, GEnSRF was spun up for 8 days prior to June 1, 2008. The lag window for the smoother was set to 10 days to take into account that CO_2 information is preserved over the continent for a maximum of 10 days. Note that a much longer lag window would have been required for global applications where there is no finite residence time for an air mass in the domain, or if substantial flux temporal correlation had been assumed *a priori*. Even with a 10-day window longer integrations of the transport model are required as well as more parameters need to be estimated. Thus to represent the covariance matrix properly, it becomes necessary to have a large number of ensemble members.

6.3.3 Evaluating the analysis

The posterior flux estimates from the different GEnSRF runs and GIM are compared using both quantitative and qualitative metrics at different spatial and temporal scales. Results

***Table 6.1** – Summary of the GEnSRF configurations reported in Section 6.4.*

Test Case Name	Inversion setup	Number of observations	Parameters		
			Ensemble size	Localization length	Prior inflation standard deviation
GEnSRF TC1E500	TC1	3-hourly with data gaps [a,b]	500	n/a	n/a
GEnSRF TC1E2500	TC1		2500	n/a	n/a
GEnSRF TC1E500_L1500I005	TC1		500	1500 km	0.05
GEnSRF TC2E500	TC2	3-hourly [a]	500	n/a	n/a
GEnSRF TC2E500_L1500I005	TC2		500	1500 km	0.05

[a.] 3-hourly implies observations are available at 8 time-periods during the day
[b.] For the shorter towers (height ≤ 150 m) afternoon measurements are only used; for the very short towers (height ≤ 30 m) only those measurements recommended by the data providers are used.

are presented for a subset of GEnSRF runs (Table 6.1) that answer the specific questions posed in the study, and other setups are discussed where appropriate.

In terms of time-averaged diagnostics, the two quantitative metrics used are the root mean square difference (RMSD) and the correlation coefficients (CC). The GEnSRF and the GIM flux estimates are aggregated to a monthly timescale and the RMSD and the CC calculated at the native 1°×1° spatial resolution for all grid-cells across the continent. Both these quantities are reported aggregated over North America (NA) and the Mid-Continent Intensive (MCI) region. The MCI region (e.g. *Lauvaux et al*. [2012]) that is shown as the green shaded area in Figure 6.2, was not only well constrained by a dense measurement network in 2008, but also lies in the interior of the study domain and hence is immune to biases that may arise along the boundaries of the study domain (*Dirren et al*. [2007]). The monthly fluxes and uncertainties are also aggregated to seven ecoregions (Figure 6.3) that are loosely defined based on the work of *Olson* (2001) and demarcate large (mostly contiguous) regions with similar climate, land cover and land use.

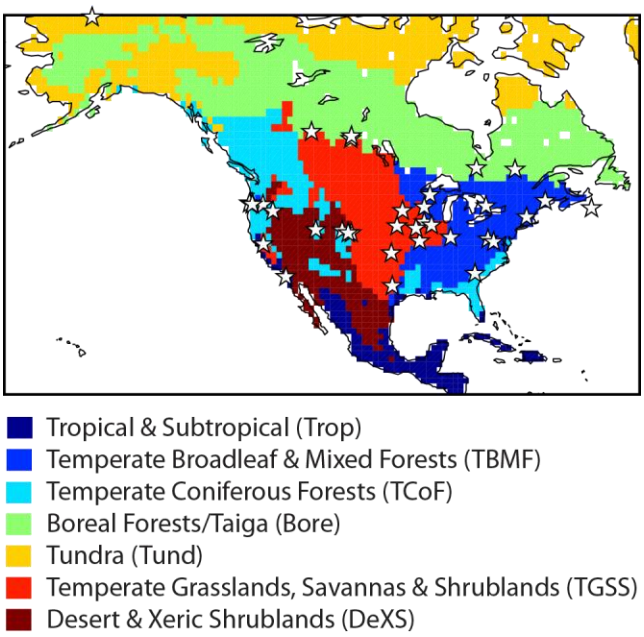


Figure 6.3 *– Ecoregion map, modified from Olson (2001), which is used for analyzing inversion results at spatially-aggregated scales. Stars represent the location of the 35 tower network.*

In terms of diagnostics at fine time scales, the GEnSRF performance is evaluated at 3-hourly and daily time scales, aggregated spatially to the full North American domain. By domain-averaging the recovered 3-hourly fluxes, we assess the ability of GEnSRF to accurately recover the diurnal cycle of the CO_2 fluxes. Daily RMSD between the GEnSRF and the GIM grid-scale flux estimates are also examined as a function of time to evaluate the filter stability.

The degree to which the GEnSRF fluxes reproduced the atmospheric CO_2 observations was also evaluated (results not shown), but a direct comparison of GEnSRF fluxes rather than atmospheric concentrations provides a more direct measure of the impact of the numerical DA scheme.

6.4 RESULTS

6.4.1 Multi-scale evaluation of the ensemble estimates for the control run (TC1E500)

Monthly-averaged grid-scale flux estimates and uncertainties for TC1 are presented in Figure 6.4. Qualitatively, it is clear that the control run (GEnSRF TC1E500 – Figure 6.4C, D) is not capable of reproducing the monthly-averaged GIM estimates or their associated uncertainties (Figure 6.4A, B). The under-estimated uncertainties should not be interpreted as more confident estimates, but rather point to the problem of insufficient variance in the ensemble. While the ensemble approach correctly captures the flux estimates over the Eastern corridor and the Southern parts of the continent, its performance degrades over the Northwestern region in Alaska and Canada, where scattered sources are incorrectly inferred throughout. In real-data application, this could be caused by two reasons: 1) the use of a limited number of ensemble members resulting in large sampling error, and 2) in general, this area has a sparse network with several of the available towers located in complex terrains where the transport is difficult to model. Given that this synthetic data study does not include transport model errors, the erroneous fluxes suggested by GEnSRF are a product of spurious ensemble noise. As indicated in Table 6.2, the difference in the spatial patterns between the two sets of estimates manifests itself in low CC and high RMSD between GIM and GEnSRF over North America.

Monthly-averaged ecoregion-scale flux estimates and associated uncertainties are presented in Figure 6.5. GEnSRF TC1E500 estimates suggest a smaller sink throughout all ecoregions relative to the GIM TC1 estimates, and the 95% uncertainty bounds based on the ensemble estimate only capture the true fluxes in 4 of the 7 ecoregions. At the continental scale, the GEnSRF TC1E500 estimate (-23.8 (±3.4) gC/(m^2month)) is significantly higher than the GIM TC1 estimate (-32.8 (±2.7) gC/(m^2month)), and unlike the GIM estimate, and does not

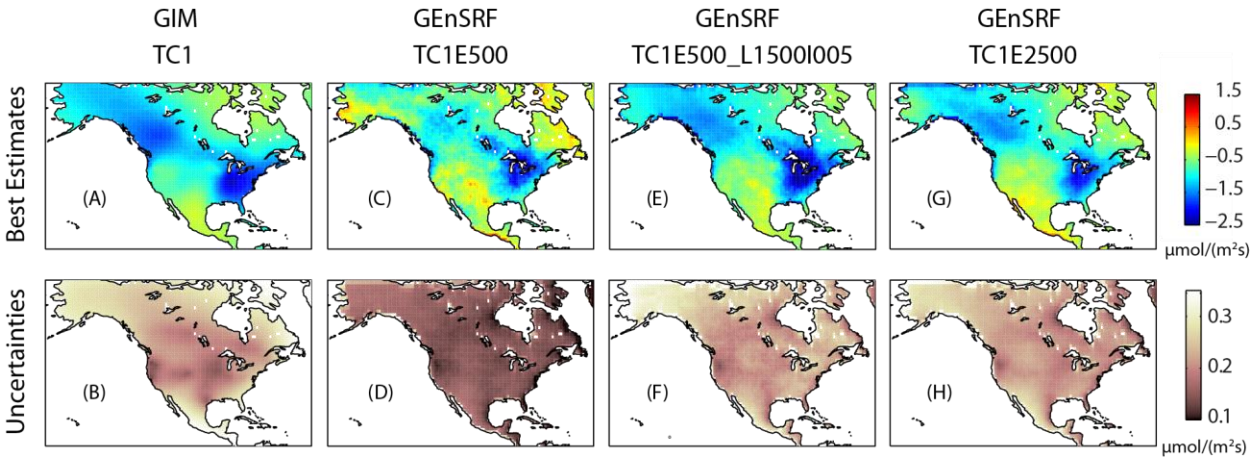


Figure 6.4 *–TC1 flux estimates (top row) and associated uncertainties (bottom row) aggregated to the monthly scale for GIM (A and B) and three different GEnSRF runs (C to H).*

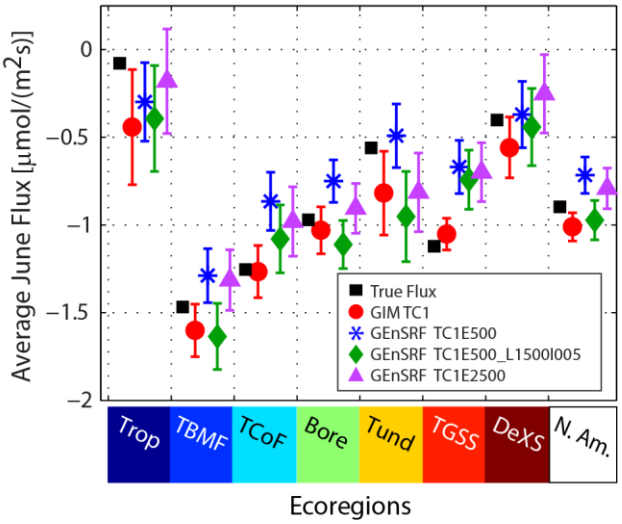


Figure 6.5 *– Estimated monthly-averaged flux estimates and the associated uncertainties aggregated to ecoregions (Figure 6.3) and continental scales. The error bars represent 95% uncertainty bounds.*

capture the true flux of -30.46gC/(m^{2}month).

The inferred monthly-averaged diurnal cycle at the continental scale is shown in Figure 6.6. GEnSRF TC1E500 does not reproduce the GIM TC1 diurnal pattern, with the difference between the two estimates spiking around 0400h and 1900h UTC. The largest differences between the estimated diurnal cycles coincide with times with the greatest temporal gradient in

the true underlying fluxes, as well as times when observation locations are coming into/out-of the TC1 network. Another mechanism that could cause these observed errors is the sampling error due to a small ensemble size which could result in spurious temporal correlations in the estimates, leading to a dampened diurnal cycle relative to GIM. Further analysis (Sections 6.4.2 and 6.4.3) suggests that the gradient in the true diurnal cycle is the better explanatory factor. Conclusions based on the inferred diurnal cycle for the MCI region, which spans a much narrower longitudinal range and therefore exhibits less smearing of the diurnal cycle, are consistent with those for the full continent (results not shown).

Overall, the conclusion from the control run is that the small ensemble size and limited observational information in TC1E500 hinder the ensemble's ability to reproduce GIM estimates across spatial and temporal scales. Sampling errors and sparse measurements may both result in a dramatic failure of the ensemble filter to infer fluxes.

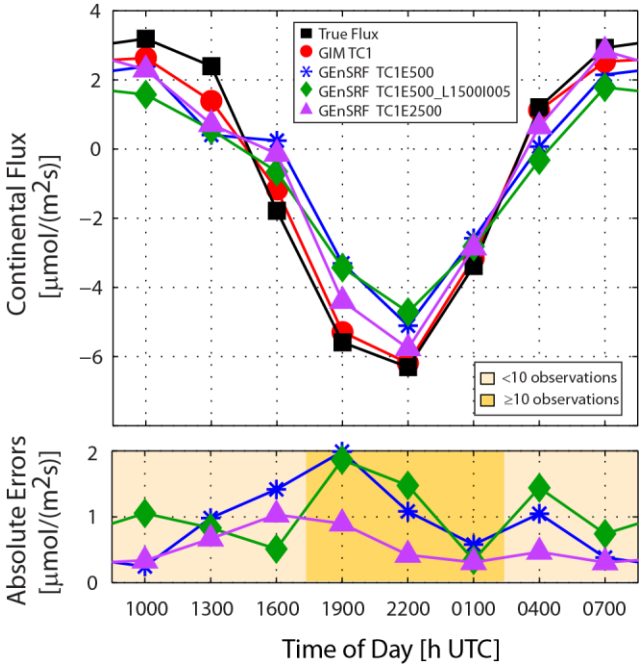


Figure 6.6 *– Estimated flux diurnal cycle (top row), and absolute errors of the individual GEnSRF estimates with respect to the GIM estimates (bottom row), aggregated to the continental scale. Also highlighted in the bottom row is the average observation density (light yellow denotes <10 observations, medium yellow denotes ≥10 observations) used in TC1 over the day.*

6.4.2 Sensitivity to ensemble size, and covariance localization and inflation algorithms

A straightforward solution to reducing the sampling error is to increase the ensemble size, which in effect increases the rank of the ensemble estimate of the prior error covariance matrix. In the absence of a dynamical model and at the limit of a large ensemble, the ensemble filter asymptotically approaches the Kalman filter (assuming the error characteristics remain Gaussian) at a convergence rate of $1/\sqrt{N}$. A large ensemble (GEnSRF TC1E2500, Figure 6.4G, H) indeed appropriately reduces the spurious noise in the best estimates at fine spatial scales, and yields uncertainty estimates close to those from GIM TC1, albeit at the expense of an increase in computational cost compared to GEnSRF TC1E500 proportional to the increase in the size of the ensemble.

An alternate approach that does not carry substantial additional computational cost is to implement covariance localization and inflation, which dampen the sampling error and improve estimates of the flux uncertainties, as seen in GEnSRF TC1E500_L1500I005 (Figure 6.4E, F). The improved performance resulting from increasing the ensemble size and implementing localization and inflation is confirmed in Table 6.2, where GEnSRF TC1E2500 and GEnSRF TC1E500_L1500I005 both show higher CC and lower RMSD values relative to the control run. Both approaches also improve the ecoregion and continental scale results (Figure 6.5). The continental scale flux estimate for both GEnSRF TC1E2500 (-28.6 (±3.9) gC/(m^2month)) and GEnSRF TC1E500_L1500I005 (-33.3 (±3.7) gC/(m^2month)) capture the true CASA fluxes within their 95% confidence intervals.

The impact of increasing the ensemble size or of implementing inflation and localization is less conclusive for the estimation of the diurnal cycle either over the full continent (Figure 6.6) or over the MCI region (results not shown). From Figure 6.6, GEnSRF TC1E2500 captures the

Table 6.2- *Correlation coefficients (CC) and Root Mean Square Difference (RMSD; $\mu molm^{-2}s^{-1}$), calculated based on grid-scale, monthly averaged flux estimates between the various runs of GEnSRF and GIM TC1 (first three rows shaded light gray) and GIM TC2 (last two rows shaded dark gray). The best results are obtained by specifying a large ensemble size, albeit at the cost of high computational cost. The bold font indicates the control run.*

Test Case	North America (NA)		Mid-Continent Intensive (MCI)	
	CC	RMSD	CC	RMSD
GEnSRF TC1E500	0.64	0.52	0.77	0.32
GEnSRF TC1E2500	0.81	0.35	0.91	0.29
GEnSRF TC1E500_L1500I005	0.75	0.37	0.83	0.35
GEnSRF TC2E500	0.68	0.48	0.77	0.30
GEnSRF TC2E500_L1500I005	0.76	0.39	0.85	0.32

diurnal cycle very well initially but the error peaks around 1600h UTC. The implementation of inflation and localization in GEnSRF TC1E500_L1500I005 does not yield a clear reduction in errors especially at 1900h UTC, although the discontinuity observed in GEnSRF TC1E500 between 1300h UTC and 1600h UTC is eliminated. The overall diurnal cycle, however, becomes even more washed out and fails to capture the true amplitude of the fluxes. The lack of error reduction resulting from the implementation of inflation and localization highlights the fact that although sampling error does contribute somewhat to the errors in the estimated diurnal cycle, the dominant cause is either the variable measurement network prescribed in TC1 or the inability of a small ensemble to capture sharp gradients in the flux diurnal cycle, or a combination of both. This is problematic, especially if we were to use these estimates either for mechanistic understanding of the carbon cycle at sub-diurnal scales, or for reconciling with estimates from biospheric models.

Changing the localization length scale to either 500 km or 3000 km for the 500-member ensemble negatively impacts the estimates (Table 6.3). A tight isotropic localization scale

***Table 6.3**- Correlation coefficients (CC) and Root Mean Square Difference (RMSD; $\mu molm^{-2}s^{-1}$), calculated based on grid-scale, monthly averaged flux estimates between the different runs of GEnSRF and GIM for TC1. The cases presented for GEnSRF specifically show the impact of localization (highlighted in light gray) and adaptive inflation (highlighted in dark gray) on the final estimates. The control run is shown in bold.*

Test Case	North America (NA)		Mid-Continent Intensive (MCI)	
	CC	RMSD	CC	RMSD
GEnSRF TC1E500	0.64	0.52	0.77	0.32
GEnSRF TC1E500_L500	0.62	0.42	0.90	0.31
GEnSRF TC1E500_L1500	0.75	0.37	0.84	0.32
GEnSRF TC1E500_L3000	0.69	0.44	0.82	0.28
GEnSRF TC1E500_L1500I001	0.75	0.37	0.84	0.33
GEnSRF TC1E500_L1500I005	0.75	0.37	0.83	0.35
GEnSRF TC1E500_L1500I025	0.75	0.38	0.79	0.44

(GEnSRF TC1E500_L500) imposes high locality, as a consequence of which the autocorrelation information modeled in the prior covariance is completely lost. Measurements impact fluxes in their immediate vicinity, while areas in which no local observations are available are not constrained at all. It is possible though that strong localization could be imposed if a wide network of measurements were available to compensate for the loss of remote influence. On the other hand, a large localization scale (GEnSRF TC1E500_L3000) cannot significantly reduce the spurious correlations among distant flux locations. This suggests that the optimal value of the localization length scale (1500 km) may be linked with the correlation length scale of the fluxes themselves (~ 1600 km, see Section 6.3.2). However, tests also revealed that the optimal filter length scale is a function of the size of the ensemble, with a smaller ensemble size requiring a shorter optimal length scale. This is due to the fact that if the number of ensemble members is large, the noise in the covariance estimates does not overwhelm the signal until much farther from the observations. This makes it harder to identify a universal mathematical or physical basis

for selecting these length scales. Nevertheless, the correlation length scale of the fluxes can be used as a starting point for the localization length scale in future filter designs.

Likewise, the estimates are found to be sensitive to the parameters of the adaptive inflation technique, especially in terms of the recovered uncertainties over data-sparse regions and periods. As evident in Table 6.3, the change in CC and RMSD is small for the different starting parameters of the adaptive inflation, but the impact is more visible when the uncertainties associated with the GEnSRF are compared to the uncertainties from the GIM. For example, with very tightly constrained inflation factors (I001), GEnSRF underestimates the standard deviation of the individual flux estimates by an average of 13% relative to GIM. Conversely, with very loose prior inflation factors (I025) the initial inflation in the ensemble is large. During assimilation of subsequent observations, the ensemble should be deflated gradually. Yet for TC1, even after the full analysis, the ensemble remains over-inflated, resulting in an overestimate of the posterior standard deviations by GEnSRF by 31%. A prior inflation factor standard deviation of 0.05 provides a good balance, with uncertainties being underestimated by GEnSRF by only 4% (Figure 6.4F).

In understanding the response of the adaptive inflation technique, two factors need to be considered: 1) the specification of a large and spatially uniform prior inflation factor uncertainty, i.e., one that does not vary between data sparse and data dense regions, and 2) a delayed response on the part of the adaptive inflation technique in adjusting to the changes in the measurement network as specific measurement location come into and out of the network throughout the day. Recall that the adaptive inflation technique is based on a Bayesian inverse modeling framework; hence, its dependency on the measurement network is not surprising. Significant improvement in the performance is obtained if the inflation is damped towards 1 as a function of time. Damping

the inflation value over time makes the technique less dependent on the measurement coverage, and has been successfully implemented in other operational tests of the adaptive technique (e.g. *Torn* [2010]).

6.4.3 Sensitivity to the measurement network

For any inversion framework based on Bayesian estimation theory, the addition of measurements in space and time will improve both the estimation accuracy and the uncertainty. As expected by increasing the temporal density of measurements in TC2, the performance of both GIM and GEnSRF estimates at the grid and ecoregion scale improve significantly (results not shown). The continental scale flux estimates for both GIM TC2 (-31.7(±1.9) gC/(m^2month)) and GEnSRF TC2E500_ L1500I005 (-32.1 (±2.7) gC/(m^2month)) improve substantially, allowing them to capture the true CASA flux estimate (-30.5gC/(m^2month)) within their 95% confidence intervals.

Of greater interest is that the GEnSRF estimates now capture the amplitude and phase of the diurnal cycle better than in the case of TC1 (Figure 6.7), even without needing to increase the ensemble size. GEnSRF TC2E500_L1500I005 estimates mirror the GIM TC2 3-hourly estimates, indicating the positive impact that the additional measurements have had on the ensemble filter, especially between 0100h to 1600h UTC. Comparing the error in the two bottom panels in Figure 6.7, one can see that the denser homogeneous network in TC2 plays a significant role in aiding the ensemble filter to correctly capture the diurnal cycle. However, the higher errors at 1900h UTC still persist, showing that these errors are more likely to be attributable to the sharp gradient in the true diurnal cycle at this time, rather than due to temporal heterogeneity in the measurement network. Either hypothesis could have been supported by results from TC1, because transition times in network size coincide with times with sharp

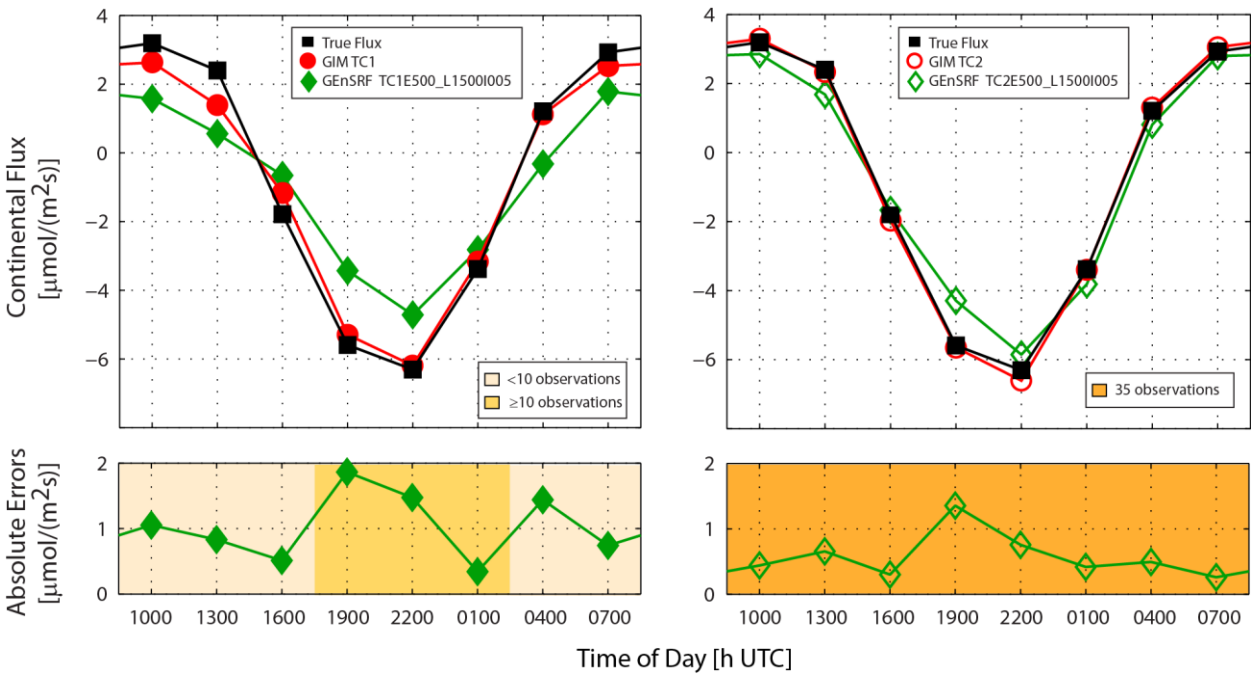


Figure 6.7 *– Estimated flux diurnal cycle (top row), and absolute errors of GEnSRF TC1E500_L1500I005 and GEnSRF TC2E500_L1500I005 with respect to the corresponding GIM estimates (bottom row), aggregated to the continental scale. For TC2 the average observation density is 35 (dark yellow) but for TC1 it varies (light yellow denotes <10 observations, medium yellow denotes ≥ 10 observations) over the day.*

gradients in the diurnal cycle. Allowing the GEnSRF to directly estimate sub-continental spatial and sub-daily temporal patterns therefore also made it possible to identify the filter sensitivity to the measurement network prescribed in TC1 and TC2.

Results from TC2 also confirm that ensemble filter performance improves with a denser measurement network. This follows from the hypothesis stated earlier in Section 6.3.2, that for an under-determined inversion problem, the ensemble system is sensitive to the spatiotemporal density of the measurements. Additional runs with a temporally-homogeneous 10 tower network confirmed that the total number of observations is a better determinant of ensemble performance at fine temporal scales relative to their temporal heterogeneity/homogeneity. Without the guidance of a dynamical model and in the absence of a rich observational constraint, the ensemble deviates from the truth, resulting in increased ensemble degeneracy and inaccurate

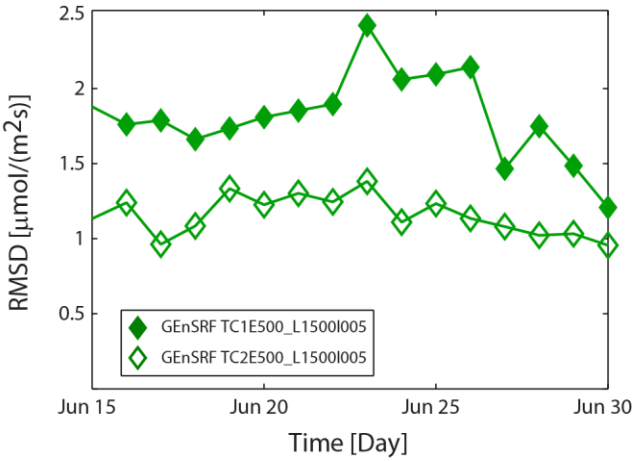


Figure 6.8 *– Time series of the Root Mean Square Difference (RMSD) between grid-scale daily-averaged estimates from GEnSRF and GIM over North America for TC1 and TC2. The time series shown here is for the latter half of the assimilation cycle to emphasize that with the measurement network in TC1, the ensemble filter does not stabilize and suffers from divergence.*

estimates. In fact, as shown in Figure 6.8, it is only in the case of TC2 that the filter is stable and reaches an asymptotic level of accuracy.

The influence of the measurement density on the ensemble behavior can be examined using several diagnostics that are commonly available in the NWP literature. All of these diagnostics, however, require knowledge of the true state against which the ensemble mean is evaluated. In this study, the true state is available from the CASA-GFED v2 fluxes, but in real application this would be unknown and hence these diagnostics could not be calculated. The diagnostic selected here examines the ratio of the time-averaged ensemble spread to the error in the ensemble mean (*Liu et al*. [2008]) at every estimation grid-point, and highlights how measurement availability controls ensemble behavior. In this case, the ensemble spread is obtained as the difference between individual ensemble members and the ensemble mean, while the error in the ensemble mean is calculated as the mean squared difference from the true state. This ratio is an indication of the optimality of the DA system, and illustrates the impact of the measurements in adjusting this ratio. Figure 6.9 shows this ratio for GEnSRF

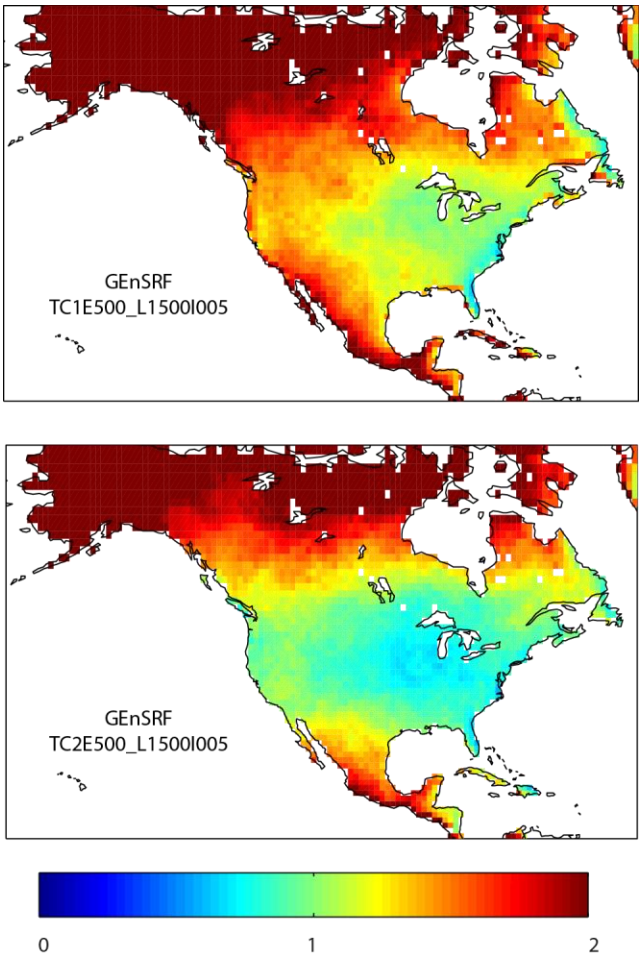


Figure 6.9 *–The ratio of grid-scale time-averaged ensemble spread and ensemble mean error for TC1E500_L1500I005 and TC2E500_L1500I005. A ratio of 1.0 (or green color) indicates optimal data assimilation.*

TC1E500_L1500I005 and GEnSRF TC2E500_L1500I005. In the case of GEnSRF TC2E500_L1500I005, the ratio is close to 1.0 over most of the continent, such that on an average the analysis spread among the ensemble members is consistent with the true errors, i.e. mean squared difference between the ensemble mean and the truth. In the case of GEnSRF TC1E500_L1500I005, however, the ratio is close to 1.0 only over a small portion of the continent that is both removed from domain boundaries and that is relatively well constrained by observations, while in other areas the value of the ratio is near 2.0. This indicates that the ensemble overestimates the uncertainties in these regions by a factor of 2. This result is

indicative of the better performance of the adaptive inflation technique with a richer observational constraint, as seen by the fact that TC2E500_L1500I005 has a reduced mismatch between the ensemble mean error and the ensemble spread, except over very sparsely observed areas like the Tundra. The dependency of the adaptive inflation technique on the spatial heterogeneity of the measurement network might seem a disadvantage at first. However, we argue that the adaptive algorithm provides inflation values that are preferable than having to manually tune the system with a single inflation value that would be applied everywhere. Figure 6.10 shows the spatially-dependent monthly-averaged inflation factors and their uncertainties, as determined by the adaptive inflation algorithm for GEnSRF TC2E500_L1500I005. Although the time-averaged values in Figure 6.10 mask the significant temporal variations of the inflation factors, they do highlight the spatial structure that is clearly consistent with the spatial density of the measurement network. This spatial variability in the inflation factors underscores the need for adopting an inflation strategy that can be adjusted recursively. Conversely, if a single inflation value had been used over the entire continent, for example a value of 1.1, then this would have under-inflated the ensemble in the data-dense regions, but over-inflated it over the data-sparse regions. This would have led to additional errors in the final estimated flux estimates and their uncertainties. Overall, results indicate that the density of the measurement network not only controls the estimation accuracy, but also ensures that the entire ensemble system and its associated algorithms function well. In the absence of a dynamical model, the measurements play an even more integral role in the assimilation process, as they drive both the ensemble mean and the ensemble spread. In practice, obtaining such a network not only requires additions to the existing monitoring network, but also improvements to atmospheric transport model that would enable the use of observations collected throughout the day, as was done in TC2.

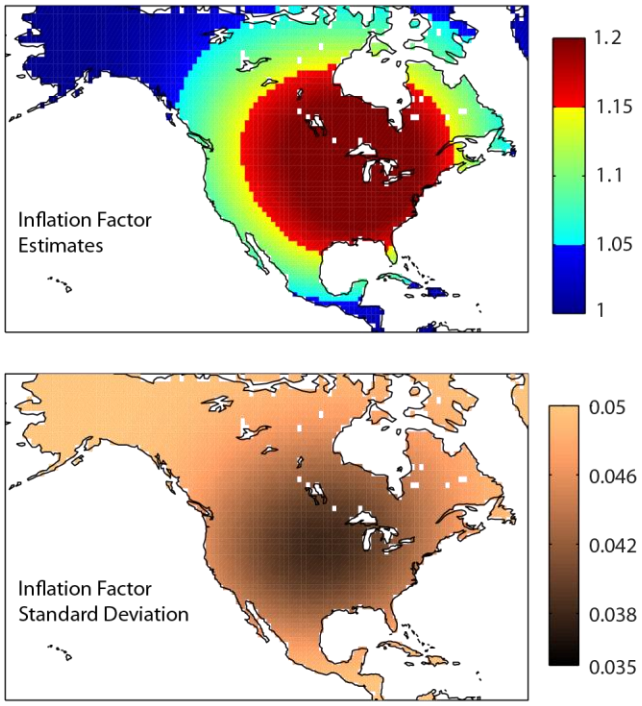


Figure 6.10 *– Monthly-averaged a posteriori inflation factor estimates and associated standard deviations for the case GEnSRF TC2E500_L1500I005. Note that the largest change in the inflation factors and the largest reduction in the prior inflation standard deviations are over areas with more measurements.*

An important question that has not been addressed here is the response of the ensemble to either transport error or measurements of varying quality. Given the sensitivity of the ensemble to the density of measurements, we expect differential and/or correlated model-data mismatch errors to play a key role as well. Tests with higher model-data mismatch covariance errors demonstrated that no additional degradation in the performance of GEnSRF was observed relative to that for GIM (results not shown). These tests did not include correlated observational errors, however.

6.5 SUMMARY

Application of data assimilation techniques for estimating sources and sinks of CO_2 provides unique opportunities to better understand the mechanistic processes governing the

carbon cycle. In this work, we examined the parameter space of the ensemble filter in terms of estimating CO_2 fluxes at high spatial and temporal resolutions. A new ensemble square root filter (GEnSRF) based on the geostatistical inverse modeling technique was presented and applied to a synthetic data study over North America.

The application of GEnSRF to different inversion regimes illustrates a dynamic interplay between three factors: 1) the spatial and temporal density of the measurements available to inform the filter, 2) the ensemble size, and the resultant sampling error, and 3) the implementation of covariance inflation and localization algorithms to ameliorate the latter. Together, these factors determine not only the relative precision and accuracy of the best estimates but also their associated uncertainties. For the ensemble filter to serve as an appropriate replacement for batch estimation of fine-scale fluxes, experiments in this study demonstrate that it may be necessary to have a dense network of measurements in space and time. To some extent, this bodes well for future applications with high-density remote sensing measurements of CO_2. Additional studies will be necessary, however, to quantify the impact of biases, correlated errors, temporal heterogeneity etc. in the remote sensing measurements.

It can be argued that the requirement for more measurements may be relaxed if a dynamical model is developed to propagate the CO_2 fluxes in time. Alternately, if the inversion can be formulated as an over-determined problem it will be better constrained by the measurements, which would also lead to better ensemble behavior. This may be problematic, however, since by solving at large spatial and temporal scales existing deficiencies in the ensemble filter are masked, and aggregation errors grow. In the long run, solving at large spatial and temporal scales may limit methodological advancements in the design of future filters for the CO_2 source-sink estimation problem.

As the popularity of the ensemble filter within the carbon science community rises, future developments will most likely revolve around reducing the impact of sampling error. In this study, this was the largest source of error resulting from the use of a limited ensemble size. Sensitivity tests with different ensemble sizes established that approximately 500 ensemble members, used in combination with covariance inflation and localization, may be used for estimating 3-hourly fluxes over North America at a $1° \times 1°$ scale. Estimates at both native and aggregated spatial scales were reliable, as were estimates at aggregated temporal scales. Capturing the diurnal cycle of the underlying fluxes proved most difficult, even when covariance inflation and localization were used. By designing inflation and localization techniques that are more tailored or customized to the CO_2 flux estimation problem, the requisite number of ensemble members may be reduced further to increase the computational efficiency. The two algorithms implemented here are drawn from NWP-related problems. Although they perform reasonably well, questions remain over the behavior of the adaptive inflation technique in data-sparse regions, the appropriateness of existing localization techniques, etc. In spite of these shortcomings, these algorithms can be implemented with a limited ensemble size to obtain reliable posterior CO_2 flux estimates but with over-inflated uncertainties over data-sparse regions, as was done here.

Chapter 7

Role of GOSAT total column CO_2 observations for the estimation of CO_2 surface fluxes

7.1 INTRODUCTION

The absence of spatially and temporally dense measurements of atmospheric CO_2 from the ground-based monitoring network is considered one of the major limiting factors to obtaining precise and accurate knowledge of CO_2 sources and sinks (e.g. *Scholes et al*. [2009]). The sparse and spatially non-uniform network is not sufficient to constrain regional fluxes with the needed certainty as well as understand the nature, geographic distribution and temporal variability of CO_2 sources and sinks (e.g. *Miller et al.* [2007]; *Le Quéré et al*. [2009]). Flux inversion studies have typically relied on atmospheric CO_2 observations from: (a) discrete flask samples (weekly snapshot measurements) of CO_2 (e.g. *Rödenbeck et al*. [2003a]; *Gurney et al*. [2003]; *Baker et al*. [2006a]; *Mueller et al*. [2008]; *Gourdji et al*. [2008]), (b) continuous *in situ* data collected over continental sites (e.g. *Carouge et al*., [2010a]; *Schuh et al*. [2010]; *Goeckede et al*. [2010]; *Gourdji et al*. [2012]), and/or (c) vertical profiles of CO_2 measurements from research aircraft (e.g. *Gerbig et al*. [2003a; 2003b]; *Stephens et al*. [2007]; *Crevoisier et al*. [2010]; *Xueref-Remy et al*. [2011a; 2011b]; *Pickett-Heaps et al*. [2011]) or commercial airlines (e.g. *Machida et al*. [2008]; *Niwa et al*. [2011]; *Patra et al*. [2011]; *Niwa et al*. [2012]), to constrain carbon fluxes and pools. In spite of the steady expansion of the *in situ* observation network, the atmospheric

inverse problem remains severely under-determined, if it is used to constrain fluxes at fine spatial and temporal scales (e.g. *Ciais et al*. [2010b]).

A viable option for overcoming the sparse coverage of the ground-based and aircraft network is to use space-based remote-sensing measurements of atmospheric CO_2. Existing remote-sensing instruments such as the TIROS Operational Vertical Sounder (TOVS), the Atmospheric InfraRed Sounder (AIRS), the Tropospheric Emission Spectrometer (TES) and the Scanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY) have been examined for their ability to measure CO_2 concentrations (e.g. *Chédin et al*. [2003]; *Engelen et al*. [2004]; *Buchwitz et al*. [2005]; *Pagano et al*. [2011]; *Schneising et al*. [2012]). These multi-purpose sounding instruments typically have increased sensitivity to the mid and the upper troposphere, which makes it difficult to relate the measured variations in CO_2 concentrations to the spatial and temporal variations of surface CO_2 sources and sinks. Thus, it is not surprising that the surface fluxes inverted from the CO_2 observations obtained from these instruments have not provided substantial insights on the carbon cycle (e.g. *Chevallier et al*. [2005b]; *Chevallier et al*. [2009a]; *Nassar et al*. [2011]).

More recently, efforts have been invested in the development of CO_2-dedicated remote-sensing instruments, such as the Orbiting Carbon Observatory-2 (OCO-2; *Eldering et al*. [2012]), the Greenhouse Gases Observation SATellite 'IBUKI' (GOSAT; *Kuze et al*. [2009]; *Yokota et al*. [2009]) and the Active Sensing of CO_2 Emissions over Nights, Days and Seasons Mission (ASCENDS; *NRC* [2007]; *NASA* [2008]); all of these instruments are sensitive to the lower troposphere, and are expected to improve our scientific understanding of regional carbon cycle processes and budgets. GOSAT, which was successfully launched in January 2009, is the only remote-sensing mission currently operational that was specifically designed to measure

atmospheric CO_2 and CH_4 concentrations from space. GOSAT takes global soundings of reflected sunlight from which column-averaged dry air mole fractions of CO_2 (XCO_2) are retrieved. The GOSAT satellite has a Sun-synchronous orbit at an altitude of 666 km and a 3-day recurrence with the descending node around 13:00 local time (e.g. *Hamazaki et al*. [2004; 2005]). Over the 3-day orbital repeat cycle, GOSAT measures thousands of single soundings that cover the globe but the most useful observations for retrieving CO_2 concentrations are limited to areas under clear-sky conditions (e.g. *Yoshida et al*. [2011]; *O'Dell et al*. [2012]). Even then the number of clear-sky locations where the GOSAT data are available (see Section 7.2.2) far surpasses the current number (~ 200) of *in situ* monitoring locations, as registered with the World Data Centre for Greenhouse Gases (WDCGG; http://ds.data.jma.go.jp/gmd/wdcgg/).

It may still be early, however, to assess the real contribution of GOSAT to the estimation of carbon sources and sinks with reliable uncertainties. This is because the GOSAT XCO_2 retrieval algorithms are themselves maturing (e.g. *O'Dell et al*. [2012]; *Oshchepkov et al*. [2012]), and studies are underway to minimize the retrieval errors and improve the precision of the data products. Theoretical studies (e.g. *Kadygrov et al*. [2009]) using simulated GOSAT observations have claimed that GOSAT XCO_2 retrievals will be able to reduce the mean regional flux uncertainties from those based on *in situ* measurements by 50% only if the total errors in the monthly averaged column data are less than 0.8 ppm. Other synthetic data studies (e.g. *Maksyutov et al*. [2008]; *Chevallier et al*. [2009b]; *Hungershoefer et al*. [2010]; *Palmer et al*. [2011]) point out that it will be necessary to use additional information from other remote-sensing missions and/or surface networks, along with the GOSAT data, to improve our current knowledge of carbon fluxes. Since the launch of GOSAT, only one study has been published (*Takagi et al*. [2011]) that focuses on the estimation of CO_2 surface fluxes from the actual

GOSAT observations. This study used monthly-mean surface observations and monthly-averaged GOSAT XCO_2 retrievals gridded to 5°×5°, for estimating monthly fluxes over 64 regions. The conclusions of the study were rather disappointing, as sound estimates of surface fluxes could not be obtained due to retrieval biases in the data product, and large aggregation errors caused by the estimation of fluxes at coarse spatial and temporal scales.

Nonetheless, it is important to keep probing the utility of the GOSAT observations as they provide valuable feedback to the retrieval community regarding the value of their data products. To the best knowledge of the author, 6 or 7 groups worldwide (*S. Houweling, pers. comm.*) are actively pursuing flux inversions with the GOSAT data at different spatial (~ 2.5° or coarser) and temporal (~ weekly or coarser) resolutions. Several of these studies are still exploratory, and a variety of Bayesian data assimilation (DA) techniques (ensemble filter or variational) are being used to extract information from the GOSAT retrievals. The use of Bayesian DA techniques limits our understanding of the actual information content of the GOSAT data as these DA systems (e.g. *Chevallier et al*. [2005b]; *Peters et al*. [2005]; *Rödenbeck* [2005]; *Baker et al*. [2006a]; *Feng et al*. [2009]; *Miyazaki et al*. [2011]) require an initial estimate of CO_2 fluxes from bottom-up models and/or inventories, or combination thereof. Hence, the posterior flux estimates are no longer a direct reflection of the information contained in the GOSAT data.

The goal of this study is twofold: (1) to assess the ability of the GOSAT XCO_2 data to constrain fine scale (*spatial* - 1° × 1.25° and *temporal* - daily) estimates of carbon flux over different seasons, and (2) to quantify the influence of the GOSAT XCO_2 observations in constraining surface fluxes, relative to the observations from the surface network. Flux estimates are presented for two representative months that are typical of Northern Hemisphere early

summer (June 2009) and winter (January 2010), in order to identify whether the seasonal variability in the fluxes is adequately captured by the GOSAT observations. These estimates of surface flux represent the total flux, including terrestrial, oceanic, and anthropogenic contributions. As pointed out in *Mueller et al.* [2008], directly estimating the total flux avoids the possibility of aliasing the uncertainties and seasonality of fossil fuel emissions (e.g. *Gurney et al*. [2003]) onto the estimated biospheric flux signal. The estimated fluxes are subsequently compared at various spatial and temporal scales to bottom-up estimates of biospheric (*Potter et al*. [2007]), oceanic (*Takahashi et al*. [2002]), fossil fuel (e.g. *Friedlingstein et al*. [2010]), and fire emissions (*van der Werf et al*. [2006]) fluxes.

The geostatistical ensemble square root filter (GEnSRF) developed and tested in the previous chapter (e.g. *Chatterjee et al*. [in press]) is well-suited to obtaining the surface flux estimates as it has been designed to: (a) handle large observational datasets, and (b) provide data-driven estimates of surface fluxes at high spatial and temporal resolutions. By reducing the influence of prior bottom-up flux estimates from biospheric models and/or inventories, GEnSRF allows us to directly assess the information content of the atmospheric observations. Further, by judiciously selecting the model of the trend as an unknown but constant set of mean fluxes in space and time, the analysis is not influenced by any auxiliary variables that may prescribe flux spatial patterns *a priori* on the surface flux estimates. Analytical tools such as the influence matrix (e.g. *Cardinali et al*. [2004]; *Liu et al*. [2009]) are used to directly assess the amount of influence each data type (i.e., GOSAT vs. surface-network) has on the posterior analysis.

The results from this study will be shared with the GOSAT-ACOS retrieval team, which was set up after the launch failure of the Orbiting Carbon Observatory (OCO) mission (e.g. *Crisp et al*. [2004]), under the auspices of the NASA Atmospheric CO_2 Observations from Space

(ACOS) task. This may provide valuable feedback to the retrieval team on the precision of the XCO_2 data product, and its ability to inform contemporary CO_2 sources and sinks. Additionally, the results will be submitted to an inter-comparison effort led by S. Houweling (GOSAT inversion inter-comparison protocol, S. Houweling, *pers. comm.*). The inter-comparison effort specifically aims to quantify the effects of GOSAT observations on estimated fluxes from several inverse models, update the global and regional carbon budgets using the GOSAT observations, and if possible, attribute the variability and trends of the CO_2 surface fluxes to the underlying drivers.

7.2 INVERSION FRAMEWORK

The geostatistical objective function used in the solution of the atmospheric CO_2 inverse problem is expressed as:

$$J(\mathbf{s},\boldsymbol{\beta}) = \frac{1}{2}\left[\mathbf{z} - h(\mathbf{s})\right]^T \mathbf{R}^{-1}\left[\mathbf{z} - h(\mathbf{s})\right] + \frac{1}{2}\left[\mathbf{s} - \mathbf{X}\boldsymbol{\beta}\right]^T \left(\mathbf{Q}^b\right)^{-1}\left[\mathbf{s} - \mathbf{X}\boldsymbol{\beta}\right] \quad 7.1$$

where $\mathbf{z}$ is a $n\times1$ vector of observations, h represents the atmospheric transport model, $\mathbf{s}$ is a $m\times1$ vector of the discretized *true but unknown* surface flux distribution, $\mathbf{R}$ is the $n\times n$ model-data mismatch covariance, $\mathbf{X}$ is an $m\times1$ vector of ones in the work but could also include auxiliary variables related to carbon flux, $\boldsymbol{\beta}$ is an unknown constant here but could also include unknown drift coefficients that scale other variables in $\mathbf{X}$, and the prior covariance matrix $\mathbf{Q}^b$ describes the expected variability in flux departures from $\mathbf{X}\boldsymbol{\beta}$ as a function of the separation distance in space and time between fluxes. Further description of these components is presented in Sections 7.2.1 to 7.2.4.

The overall mathematical framework of GEnSRF DA system that is used to solve Equation 7.1, including the implementation of the covariance localization and the adaptive

inflation algorithms, is based on *Chatterjee et al*. [in press]. The operational implementation, however, is more advanced mainly due to the data volumes that need to be assimilated. In order to avoid repetition of the entire framework, the reader is referred to Chapter 6 (Section 6.2) and the specific components of the problem that are different from Chapter 6 are discussed here.

7.2.1 Flux estimation resolution

CO_2 fluxes are estimated *daily* at $1° \times 1.25°$ for both land and ocean components. Although the choice of the spatial and temporal resolution is limited by the resolution and model integration time step of the transport model (Section 7.2.3), these are substantially finer than any current global flux inversion study using the GOSAT data. The primary reason for estimating fluxes at such a fine spatial and temporal scale is to minimize the aggregation errors. Aggregation errors occur when the atmospheric measurements are sensitive to variability in the fluxes at finer scales than the scale at which the inversion is allowed to adjust the fluxes (e.g. *Kaminski et al*. [2001]; *Peylin et al*. [2002]). Hence, we follow the recommendation of *Kaminski et al*. [2001] that global flux inversions should resolve sources and sinks at as fine a scale as computationally feasible (with the limit being the resolution of the transport model).

Ideally the temporal resolution should be sub-daily to minimize the temporal aggregation error as pointed out by several recent studies (e.g. *Gourdji et al*. [2010; 2012]; *Huntzinger et al*. [2011]; *Chatterjee et al*. [in press]). It is worth noting, however, that these aforementioned studies used continuous tower data from regions with high flux variability, which is influenced by the diurnal cycle from terrestrial ecosystems and local sources and sinks of CO_2. The surface flask sites used in the current study are primarily located over remote areas, where the sampled air shows little diurnal variation caused by local CO_2 fluxes. In terms of the satellite observations, the diurnal amplitude of column CO_2 is less than 1 ppm even over highly

productive terrestrial ecosystems (e.g. *Olsen and Randerson* [2004]; *Kawa et al*. [2010]). Hence, we believe a daily estimation scale using the remote-sensing and/or the surface flask data may cause less temporal aggregation error than it would have with continuous tower data.

7.2.2 Atmospheric CO_2 observations

Figure 7.1A and 7.1B shows a snapshot of the atmospheric CO_2 observations used in the study that are obtained from the GOSAT instrument and a ground network of surface flask sites. The GOSAT L2 retrieval data have been generated by version 2.9 of the ACOS algorithm (*O'Dell et al*. [2012]; *Crisp et al*. [2012]). Based on the recommendations of the ACOS team, only the high (H) gain observations are retained from the data product, which are then filtered through a conservative scheme based on *Wunch et al*. [2011b]. For a month, this amounts to retaining 10-15% (i.e., ~6e3) of the total observations (i.e., ~ 4e4) from the native ACOS L2 data product. Note that the GOSAT data, as used in the current study, are not bias-corrected, although first estimates for bias correction do exist (*Wunch et al*. [2011b]).

CO_2 concentration measurements are also obtained from approximately 100 surface flask sites (Figure 7.1E and 7.1F) within the NOAA Earth System Research Laboratory (ESRL) Global Monitoring Division cooperative air sampling network (*Conway et al*. [2011]). The surface flask observations are available at weekly time scales, and a typical site provides 4 to 5 observations within a month. In addition, several locations have missing data during the examined time periods, reducing the number of operational sites to ~70 per week.

7.2.3 Atmospheric tracer transport model

An off-line atmospheric transport model (GSFC parameterized chemistry and transport model – 'PCTM', *Kawa et al*. [2004]), which is run internally as part of the GEnSRF DA

system, is used to relate surface CO_2 fluxes to atmospheric CO_2 observations. PCTM is driven by pre-calculated meteorological fields (horizontal winds, surface pressure, vertical diffusion coefficient, and cloud-convective mass flux) from NASA's GEOS4-DAS reanalysis at a resolution of 1° ×1.25° in latitude/longitude and 55 vertical layers. The model uses a vertically-Lagrangian finite volume advection scheme (*Lin* [2004]) and has simple linear schemes for both dry and convective vertical mixing.

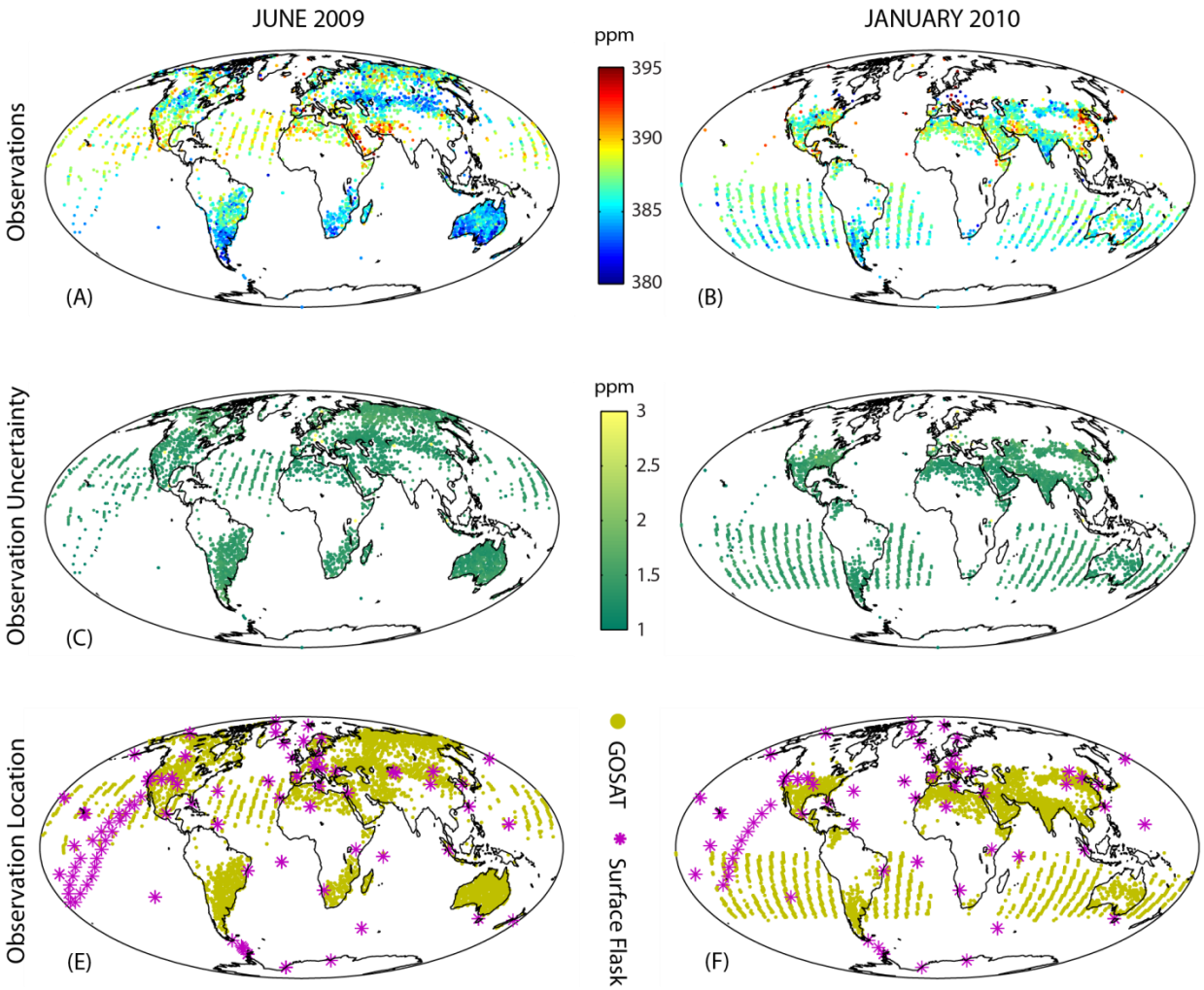


Figure 7.1 *- GOSAT ACOSv2.9 XCO_2 and surface flask CO_2 observations, their associated standard deviations and their locations, gridded to 1° × 1.25°. Note the seasonal shift in the north-south upper bounds of the GOSAT XCO_2 retrievals over the oceans between Panels E and F.*

7.2.4 Error covariance matrices

The model-data mismatch covariance matrix **R** is a diagonal matrix whose elements represent the variances associated with measurement, transport, representation and aggregation errors (*Engelen et al*. [2002]) for each observation.

For the GOSAT XCO_2 observations, the model-data mismatch variances are obtained by adding 1.0 ppm^2 to the square of the measurement errors reported with the ACOS v2.9 data product. The additional 1.0 ppm^2 (e.g. *Chevallier et al.* [2010b]) is assumed to be representative of the transport model, representation and aggregation errors in column-averaged observations. Overall, for the filtered H-gain GOSAT data the prescribed model-data mismatch uncertainties range from 1.0 to 2.0 ppm (i.e., model-data mismatch variance = 1.0 ppm^2 to 4.0 ppm^2).

For the surface flask observations, the errors are typically site-dependent (e.g. *Rödenbeck* [2005]; *Michalak et al*. [2005]; *Mueller et al*. [2008]; *Hungershoefer et al*. [2010]). Following the setup outlined in *Hungershoefer et al*. [2010], the surface flask sites are sub-divided into four groups, and the uncertainty for each group of sites specified accordingly:

- Remote sites (island, deserts, Antarctica): 1.0 ppm
- Shore sites with mixed ocean/continent influence: 1.5 ppm
- Mountain site (within continents): 1.5 ppm
- Continental site with proximity to large sources and sinks: 3.0 ppm

The model-data mismatch covariance matrix is obtained by squaring the prescribed uncertainties for each observation. Figure 7.1C and 7.1D shows a snapshot of the total errors (i.e., sum of measurement, transport, representation and aggregation errors) prescribed for all the observations used in the study.

The prior covariance matrix $\mathbf{Q}^b$ contains off-diagonal entries describing the spatial correlation of the flux deviations from the model of the trend $\mathbf{X\beta}$. Similar to Chapter 6, $\mathbf{Q}^b$ is prescribed as a block diagonal matrix, with each block describing the correlation between grid-scale fluxes for each time period of the inversion. Each block is modeled by an exponential covariance function:

$$\mathbf{Q}^b\left(d \mid \sigma^2, l\right) = \sigma^2 \exp\left(-\frac{d}{l}\right) \tag{7.2}$$

where d is the spatial separation distance between the grid points where fluxes are to be estimated, σ^2 represents the variance of the flux residuals at large separation distances, and l is the range parameter. The correlation length beyond which correlation between the flux residuals becomes negligible is defined as approximately $3l$ (*Chiles and Delfiner* [1999]) for an exponential model.

It is necessary to distinguish between the land and ocean components as different processes drive the underlying fluxes. Hence, spatial correlation is assumed among land and ocean flux residuals but not between them (e.g. *Michalak et al*. [2004]). The parameters of the exponential covariance function are based on a variogram analysis of the spatial variability of a typical bottom-up estimate, which is based on the net ecosystem production from CASA AMES (*Potter et al*. [2007]), oceanic carbon exchange from *Takahashi et al*. [2002], fossil fuel emissions extrapolated from the CDIAC database (e.g. *Friedlingstein et al*. [2010]), and fire emissions from the GFEDv2 (*van der Werf et al*. [2006]). The land variance ($\sigma^2_{\mathbf{Q}_{land}}$ = 0.45 [μmol/(m^{2}s)]^2) is almost a magnitude higher than the ocean variance ($\sigma^2_{\mathbf{Q}_{ocean}}$ = 0.020 [μmol/(m^{2}s)]^2) while the terrestrial flux correlation length ($l_{\mathbf{Q}land}$ = $3l$ = 2700 km) is slightly lower than the ocean flux correlation length ($l_{\mathbf{Q}ocean}$ = $3l$ = 3200 km). Note that the covariance

parameters specified here may change if the specified land and ocean bottom-up estimates are modified.

Rather than subjectively specifying the errors for each observation set individually (i.e., for **R**) or deriving it from a single bottom-up estimate (i.e., for $\mathbf{Q}^b$), a more robust approach is to statistically infer the covariance parameters from the atmospheric measurements directly via statistical approaches (e.g. *Michalak et al*. [2004]; *Michalak et al*. [2005]; *Mueller et al*. [2008]). These approaches, however, are substantially more computationally intensive and require the pre-calculation and storage of the transport model (i.e., generate the full sensitivity matrix **H**), which is infeasible in this case. In the future, these statistical approaches will be adapted to a data assimilation framework in order to specify the error covariance matrices directly from the atmospheric data.

7.2.5 Lag window

Peters et al. [2005] argued that the length of the lag window for a Kalman smoother must vary with the flux estimation temporal resolution. This argument was based on the fact that for monthly flux inversions, *Bruhwiler et al*. [2005] needed to link observations to 6-9 months of past fluxes, whereas for weekly flux inversions *Peters et al*. [2005] required only 8-10 weeks of lag to reliably retrieve CO_2 fluxes. An important argument that can be made here is that the length of the lag window may also be a function of the density of the observation network. Both *Bruhwiler et al*. [2005] and *Peters et al*. [2005] primarily used sparse near-surface CO_2 concentration observations whereby the large distances between observations required a long assimilation window to collect enough constraints on the fluxes on continental and hemispheric scales. But the density of the observation network used in *Peters et al*. [2005] was relatively higher than the observation density employed by *Bruhwiler et al*. [2005]. For our case, this

implies that the relatively good coverage of satellite CO_2 observations and the daily estimation scale may allow us to use a shorter lag window compared to *Bruhwiler et al*. [2005] or *Peters et al*. [2005]. Also, this will allow us to capture the instantaneous information in the surface carbon fluxes before the information in the CO_2 concentrations is blurred by the atmospheric transport.

The satellite data provides column-averaged CO_2 concentrations, which have information on surface fluxes from different times and places. *Miyazaki et al*. [2011] claimed that because of this mixed information in the column data, it is more difficult to extract meaningful information from the column-averaged data relative to surface data, even if a large lag window is used. A longer lag window would also require a larger number of ensemble members, failing which spurious temporal correlations may arise that would degrade the analysis. Hence, *Miyazaki et al*. [2011] suggested using a 3-day window, which is equivalent to the ground track repeat cycle of GOSAT. The disadvantage of such a short lag window, however, is that the observation information gets too localized in time and space. As a result, variations in the surface fluxes occurring over unobserved regions, or longer-term variations occurring over weekly or monthly time scales, cannot be captured, because the observations are allowed to inform only local (i.e., in space) and recent (i.e., in time) sources and sinks.

Keeping these issues in mind, the lag window in this study is set to be 21 days as a compromise between realistically capturing the longer-term variations and reducing the influence of spurious temporal correlations. Sensitivity tests were conducted (results not shown) in which the lag window was decreased to 10 days. The surface fluxes inferred, especially over the ocean regions, were found to be smoothed out indicating that the assimilation system was unable to estimate the gradual variations in the ocean fluxes with a short lag window. Using a 3-week lag

window, the size of the state vector or the total number of surface fluxes to be estimated at each assimilation time step, is approximately 1.1e6.

7.2.6 Parameters of the DA system

GEnSRF is run with an ensemble size of 500. Due to the limited number of ensemble members used, it is necessary to implement both covariance localization (i.e., to reduce spurious long-distance correlations caused by sampling error; also see Section 6.2.3) and covariance inflation (i.e., to avoid filter divergence by increasing the spread of the ensemble; also see Section 6.2.4). These algorithms introduce additional tuning parameters (i.e., localization length scale and prior inflation factor) into the DA system, which are determined via sensitivity tests as described below. Note that since the true fluxes are unknown, the flux estimates are only gauged qualitatively against the bottom-up estimate described in Section 7.2.4.

Three different localization length scales were tested – 1500 km, 3000 km and 4500 km, out of which the optimal localization length scale was found to be 3000 km. A shorter localization length scale (i.e., 1500 km) did not allow the dense observations over land to adjust the ocean fluxes. When the larger localization length scale (i.e., 4500 km) was specified, it failed to dampen the sampling error sufficiently, resulting in noisy estimates. These tests reinforced the conclusions from *Chatterjee et al.* [in press] that the optimal value of the localization length scale (3000 km) may be linked with the correlation length scale of the fluxes themselves, which in this study were inferred to be 2700 km for land and 3200 km for the oceans (see Section 7.2.4).

Even though the inflation algorithm is adaptive in nature, it requires *a priori* estimates of inflation factors and their associated variance. *Chatterjee et al*. [in press] noted that the prescribed standard deviation of the prior inflation factor plays an important role in defining the

posterior uncertainty of the estimates. In order to determine the optimal value of the inflation factors and their associated variance that one should prescribe, it is desirable to have the uncertainty estimates from a batch setup relative to which the GEnSRF uncertainty estimates can be gauged. Since obtaining a batch solution is infeasible in this case, future work will examine alternate ways by which the prior inflation factor and its associated variances can be prescribed. Presently, based on the recommendation of *Chatterjee et al*. [in press] a prior inflation factor of 1 with a standard deviation of 0.1 has been specified. Finally, as the adaptive inflation algorithm has a delayed response in adjusting to the changes in the observational locations, the inflation is damped towards 1 as a function of time.

7.3 EVALUATING THE FLUX INVERSIONS

For each of the two examined months (i.e., June 2009 and January 2010), flux estimates are obtained using different sets of CO_2 observations: (a) using both surface flask and GOSAT together (henceforth, termed as *GOSAT+SF*), (b) using only the GOSAT retrievals (henceforth, termed as *GOSAT-only*) and (c) using only the surface flask (henceforth, termed as *SF-only*). Estimated fluxes and the associated uncertainties are aggregated to a monthly time scale and then compared with each other, either at the grid scale or aggregated to the regions (Figure 7.2) used in the TransCom intercomparison study (e.g. *Gurney et al*. [2003]; *Baker et al*. [2006b]). Given that the *true* surface fluxes are unknown, the aggregated *a posteriori* fluxes are compared to the aggregated bottom up estimate that is used to obtain the covariance parameters (Section 7.2.4).

Note that the bottom-up estimate serves only as a reference point to evaluate the ability of the DA system to identify ‘reasonable’ regional-scale fluxes. Even though the bottom-up estimate is not the ‘truth’, it does provide a sense of the magnitude of the fluxes, and whether the magnitude is consistent with the source-sink estimate from the bottom-up model.

Throughout this study the flux estimates from the SF-only inversion are considered the baseline against which the increments made by the GOSAT observations (i.e., GOSAT + SF) are evaluated. Following *Takagi et al*. [2011], first the benefit of adding the GOSAT observations is captured in terms of a metric that measures the reduction in uncertainty (UR) relative to the *a posteriori* inversion using the surface flask network. The uncertainty reduction is obtained as a percentage, separately for each TransCom region, as:

$$\mathrm{UR} = \left(1 - \frac{\sigma_{SF+GOSAT}}{\sigma_{SF\text{-}only}}\right) \times 100 \qquad 7.3$$

where $\sigma_{SF\text{-}only}$ and $\sigma_{SF+GOSAT}$ denote the *a posteriori* uncertainty over a TransCom region from the SF-only and the GOSAT+SF inversion, respectively. These uncertainties are obtained by averaging across the estimated grid-scale uncertainties within a particular TransCom region.

Second, the observational influence on the estimates of the CO_2 surface fluxes is quantified via the information matrix (e.g. *Cardinali et al*. [2004]; *Liu et al*. [2009]). The information matrix ($\mathbf{S}^{\mathrm{o}}$) characterizes how the assimilation system uses the observations to pull the analysis away from the prior. For GEnSRF, the information matrix is mathematically given by:

$$\mathbf{S}^{\mathrm{o}} = \frac{\partial \hat{\mathbf{z}}}{\partial \mathbf{z}} = \mathbf{R}^{-1}\mathbf{H}\mathbf{Q}^{a}\mathbf{H}^{T} \qquad 7.4$$

where $\hat{\mathbf{z}}$ is projecting the final analysis estimate $\hat{\mathbf{s}}$ onto the observation space. In the ensemble setup, $\mathbf{H}\mathbf{Q}^{a}\mathbf{H}^{\mathrm{T}}$ at the end of the assimilation can be readily approximated as:

$$\mathbf{H}\mathbf{Q}^{a}\mathbf{H}^{T} \approx \frac{1}{N-1}\left(h\left(\mathbf{s}'^{a}\right)\right)\left(h\left(\mathbf{s}'^{a}\right)\right)^{T} \qquad 7.5$$

where $\mathbf{s}'^{a}$ represents the updated ensemble once all observations have been assimilated, and N is the ensemble size. From the GEnSRF setup, $h\left(\mathbf{s}'^{a}\right)$ is updated during assimilation of the

observations via Equation 6.9 (Chapter 6 – Section 6.2.2) and no extra computations are necessary to calculate this diagnostic separately. The diagonal elements of the matrix $\mathbf{S}^{o}$ ($n \times n$) are the analysis self-sensitivities and the off-diagonal elements represent the cross-sensitivities. As stated in *Cardinali et al*. [2004], the self-sensitivity with respect to the *i*th observation is $\mathbf{S}^{o}_{ii}$, while the sensitivity with respect to the prior projected to the same observational variable at the same location and time is 1- $\mathbf{S}^{o}_{ii}$. The self-sensitivity has no units and its theoretical value is between 0 and 1, when the observation errors are not correlated (i.e., **R** is diagonal). A value of 1 indicates that the analysis derives all the information from the observations, while a value of 0 indicates that all the information in the analysis comes from the prior.

The observational influence analysis provides: (a) the amount of information that the analysis extracts from the observations during data assimilation, (b) an identification of the subset of observations that are the most influential (e.g. *Lupu et al*. [2012]), and (c) a rigorous mathematical framework to choose amongst different observational datasets in the future, when multiple remote-sensing data products are to be assimilated. Note that other variations of the influence matrix can be developed using information theory (e.g. *Engelen and Stephens [2004]; Zupanski et al*. [2007b]). Although there are subtle differences in the way observation impacts

Figure 7.2*- Location of 11 land and 11 ocean TransCom regions (e.g. Gurney et al. [2003])*

are measured in these approaches, ultimately they all provide complementary information about the impact of the observations on the posterior analysis.

7.4 RESULTS

7.4.1 Monthly averaged grid-scale flux estimates

CO_2 flux estimates over land, based on both the surface flask and the GOSAT observations (i.e., GOSAT+SF) are shown in Figures 7.3A and 7.3B. Comparison with the baseline surface flask estimates (i.e., SF-only) in Figures 7.3C and 7.3D illustrate that the flux estimates using the GOSAT data infer stronger sources and sinks, especially during the Northern Hemisphere summer. In June, the flux estimates from the two experiments are noticeably different over Eurasia (carbon sinks) and Temperate Asia (carbon source). This difference is attributable to the larger number of GOSAT retrievals available over the Northern Hemisphere during June that constrains the GOSAT+SF flux inversion more than the SF-only flux inversion. Conversely, the addition of GOSAT data in January tends to make limited increments to the flux estimates from the SF-only experiment. In this case, not only are the number of GOSAT retrievals limited over land but also the quality of these retrievals is poor. Consequently, the DA system extracts less information over land from the GOSAT observations in January compared to the June observations.

Over the ocean regions (Figure 7.4) the flux estimates from the GOSAT+SF experiment shows a large modification relative to the flux estimates from the SF-only experiment. Noticeable differences are seen over the North Pacific in June and over all the ocean regions in January. Specifically, this can be attributed to the large number of GOSAT retrievals available over the ocean regions relative to the observations from the surface flask sites.

Using both the GOSAT and surface flask observations together, the global land regions are found to act as a carbon sink in June 2009 (-2.74 ± 2.33 GtC/yr, or -0.22 ± 0.19 GtC/month) with a larger tendency towards uptake in the Northern lands, but a source in January 2010 (3.02 ± 2.11 GtC/yr, or 0.25 ± 0.17 GtC/month). A possible reference for gauging these numbers is provided by the bottom-up estimate described earlier in Section 7.2.4. The bottom-up estimate provides a carbon sink estimate of -4.96 GtC/yr (i.e., -0.41 GtC/month) in June 2009 that is captured within the 95% uncertainty bound of the GOSAT+SF inversion estimates. For January, the bottom-up estimate provides a source magnitude of 11.18 GtC/yr (i.e., 0.93 GtC/month) that is significantly greater than the estimates obtained from the GOSAT+SF inversion. Recall that even though both the estimated fluxes and the bottom-up estimates report the total flux, a single

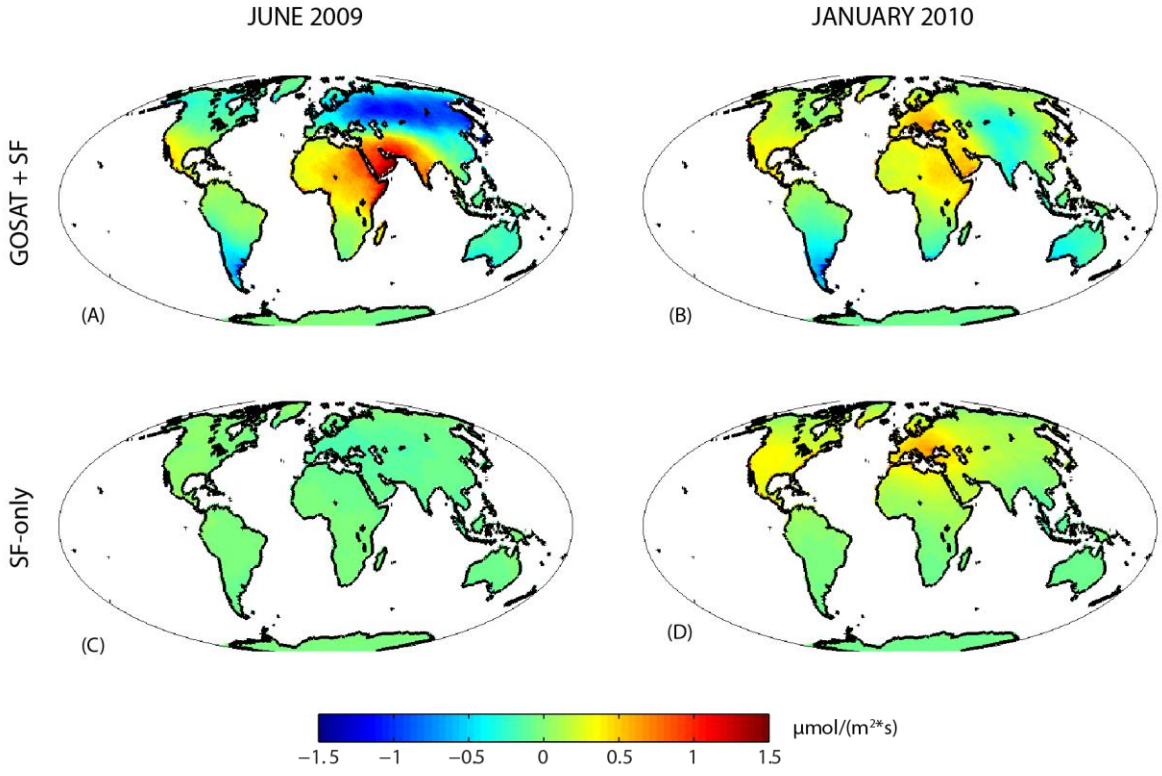


Figure 7.3*- Monthly-averaged grid-scale flux estimates over* ***land*** *for June 2009 and January 2010 using both GOSAT + SF observations (A and B), and using only SF observations (C and D).*

bottom-up estimate only serves as a qualitative reference point and cannot be used to judge the skill of the inversion. In the future a suite of bottom-up models will be used to better assess the agreement between the inversion and the bottom-up estimates (e.g. *Gourdji et al*. [2012]).

Analysis of the grid-scale flux estimates and subsequent comparison of the carbon budgets for June and January highlight a possible caveat associated with the implemented setup. Unlike other existing CO_2-DA systems, GEnSRF does not specify explicit prior information from biospheric models and/or inventories to help constrain the flux estimates. While this avoids any potential biases associated with these prior assumptions, it also fails to capture the small-scale variability in the fluxes. Within a geostatistical framework, it has been shown (e.g. *Gourdji et al*. [2008; 2012]) that small-scale flux variability can be captured if the model of the trend

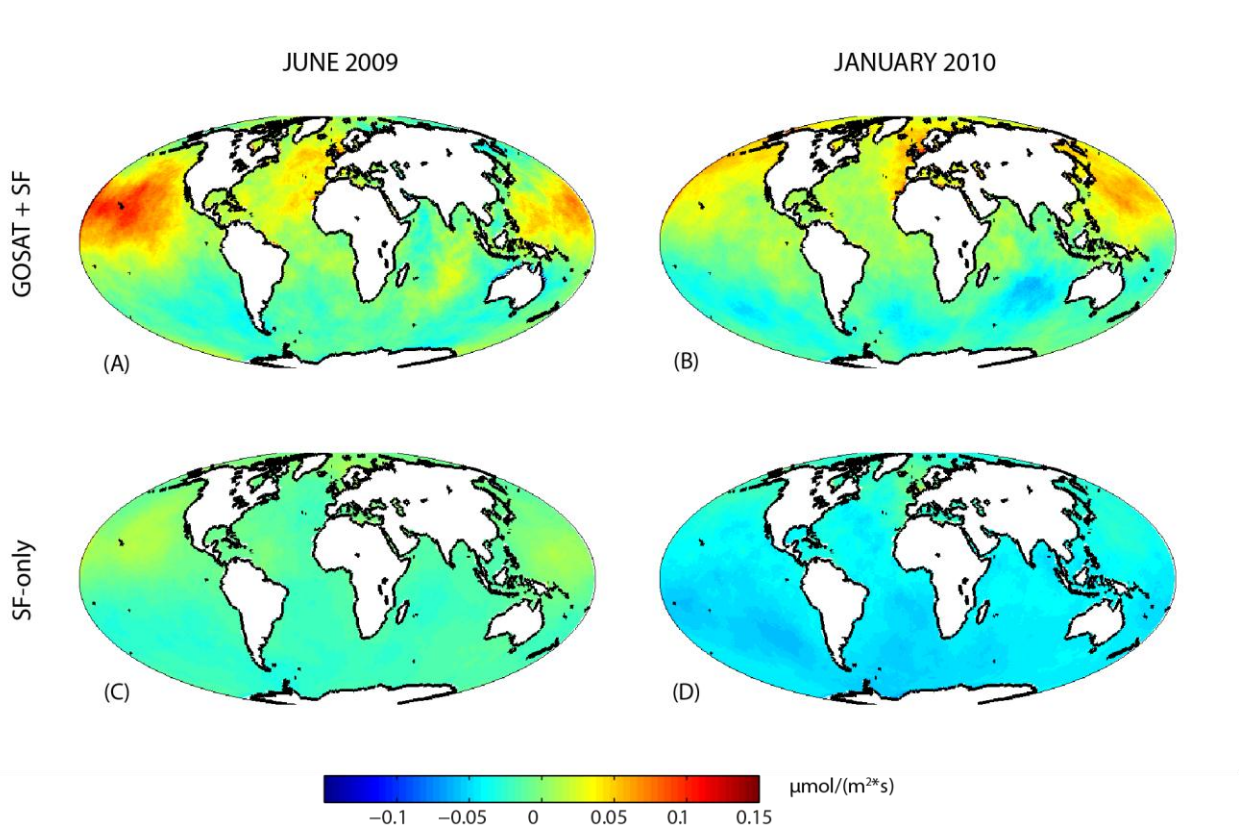


Figure 7.4*- Monthly-averaged grid-scale flux estimates over* ***ocean*** *for June 2009 and January 2010 using both GOSAT + SF observations, and using only SF observations.*

incorporates information from auxiliary environmental variables related to carbon flux. The caveat is that the column-averaged observations must be able to infer the relationship between the auxiliary variables and the fluxes, in order to select the flux covariates that optimally explain the variability in the atmospheric data. It is possible to specify all possible auxiliary environmental variables into the model of the trend. If there is a lack of atmospheric constraint from the GOSAT XCO_2 retrievals, however, then the inferred drift coefficients ($\boldsymbol{\beta}$) will not be significant and the complex model of the trend will not explain any more of the variability in the flux distribution compared to the current simple model of the trend. The complex model of the trend may even introduce spurious correlations that may potentially bias the flux estimates in under-constrained regions (e.g. *Gourdji et al*. [2012]). Thus, for the current study, the prime reason behind specifying a simple model of the mean is to compare the relative influence of the GOSAT retrievals and the surface flask observations on the analysis, without being impacted by prior spatial patterns or spurious relationships due to additional environmental variables. Obtaining realistic carbon budgets is a high priority, and future work will aim to evaluate the ability of satellite observations to select the flux covariates as well as integrate the information from the selected variables into the DA system.

7.4.2 Monthly averaged TransCom-scale flux estimates

Aggregating the flux estimates to the TransCom regions shows that when both the GOSAT and the surface flask observations are used in conjunction, the posterior analysis uncertainties are reduced over both land and ocean regions compared to the estimated uncertainties derived from the SF-only inversion (Figure 7.5). In June 2009, pronounced reductions are found in the uncertainty estimates for southern South America (15%), Boreal and Temperate Asia (12%), and over most of the ocean regions where the scarcity of the surface

flask sampling network is evident (Figure 7.1E). For January, the largest reduction in uncertainty occurs over the ocean regions (~30-35%). The lowest uncertainty reduction is found over Northern Africa (5%) for June 2009, which can be attributed to the limited number of GOSAT retrievals available over this region, the poor quality of these available retrievals, and a relative abundance of surface flask observations surrounding the Northern African region (Figure 7.1E).

Figures 7.6 and 7.7 show the flux estimates aggregated to the TransCom ocean and land regions separately. From Figure 7.6, it can be seen that when both the GOSAT and surface flask observations are used together, it not only reduces the uncertainty estimates but also provides ocean flux estimates that are more consistent with the bottom-up estimate. For example, for the Northern and Tropical ocean regions in June 2009 (Figure 7.6A), and for the Tropical and the Southern Ocean regions in January 2010 (Figure 7.6B), the flux estimates from the GOSAT+SF inversion are closer to the bottom-up estimate and have lower uncertainties associated with them. The shift from the Northern ocean regions in June to the Southern ocean regions in January corresponds to the north-south shift in the availability of good-quality retrievals from GOSAT (Figure 7.1E and7.1F). These results also demonstrate that the sparse surface flask network, especially over the Southern Atlantic and the Indian oceans, cannot constrain the fluxes over these regions unlike the high-density observations from the GOSAT. In the future, as the errors associated with the GOSAT retrievals decrease, the GOSAT observations may play a crucial role in improving our knowledge of ocean fluxes, and consequently our attempts to close the global carbon budget.

Over the land TransCom regions (Figure 7.7) the flux estimates from the GOSAT+SF inversion is able to capture the bottom-up estimate within its 95% uncertainty bounds, over the majority of the TransCom regions. For example, over Boreal Asia and Europe during June 2009,

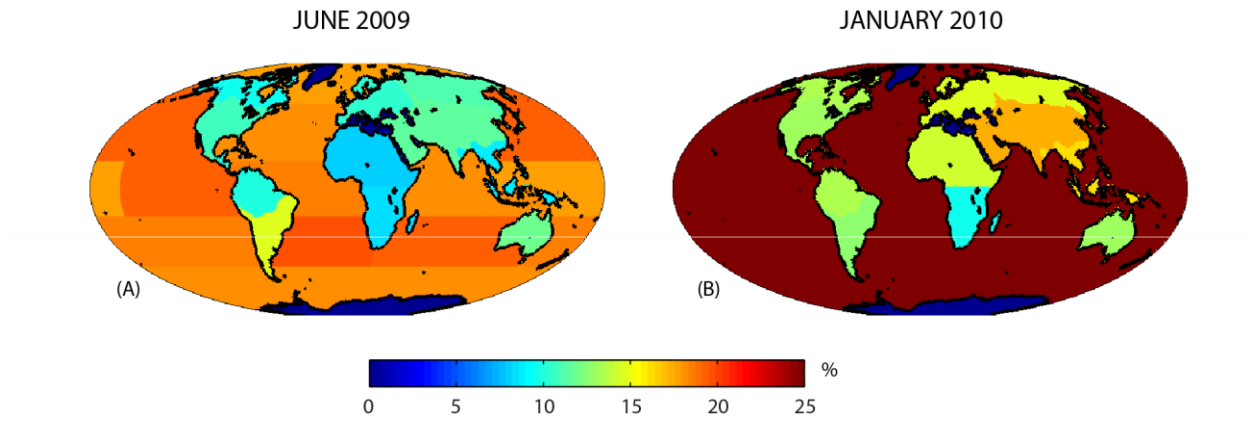


Figure 7.5 *– Reduction in the uncertainty of monthly surface CO_2 flux estimates, attained via the addition of GOSAT XCO_2 observations to the surface flask observations.*

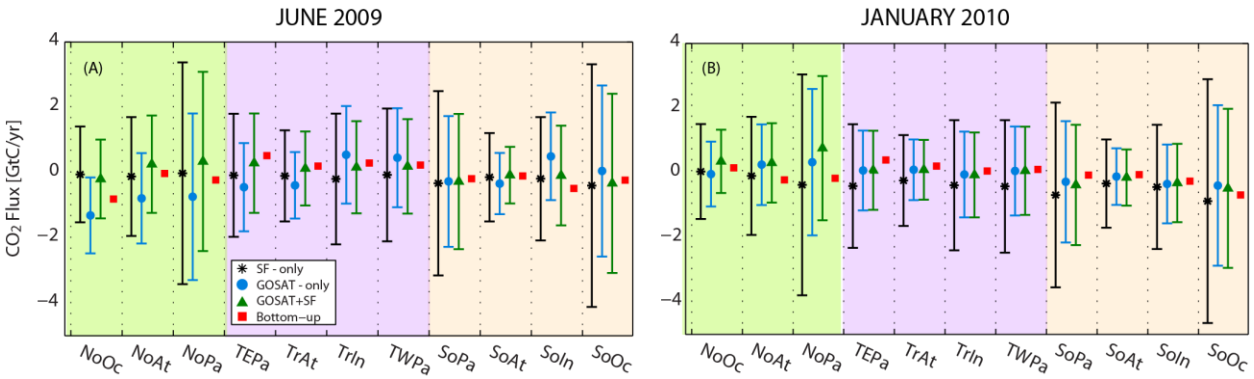


Figure 7.6 *– Estimated monthly-averaged ocean flux estimates and the associated uncertainties aggregated to the TransCom ocean regions. The error bars represent 95% uncertainty bounds. The Northern, Tropical and Southern ocean regions are shaded in green, purple, and light yellow, respectively.*

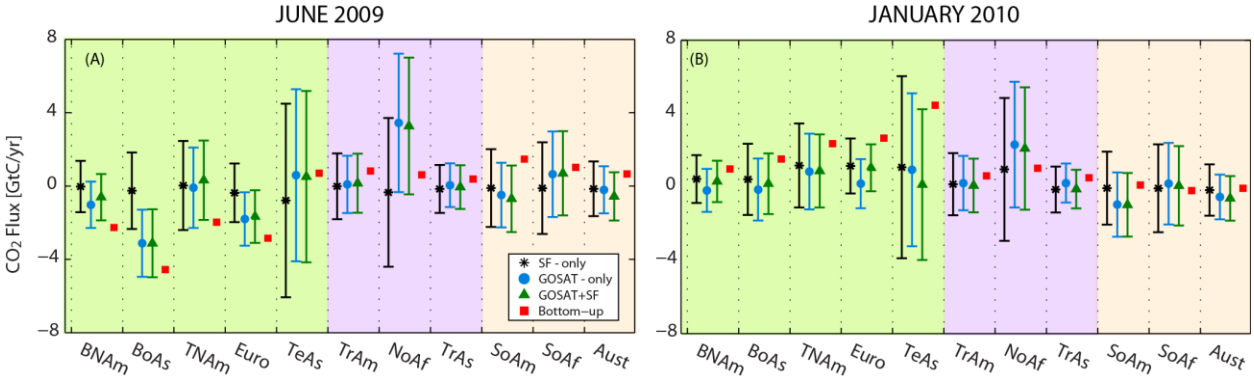


Figure 7.7 *– Estimated monthly-averaged land flux estimates and the associated uncertainties aggregated to the TransCom land regions. The error bars represent 95% uncertainty bounds. The Northern, Tropical and Southern land regions are shaded in green, purple, and light yellow, respectively.*

the sink estimate from the GOSAT+SF inversion is closer to the bottom-up flux estimate relative to the estimates from the SF-only inversion. Also the posterior analysis uncertainties are reduced over all the land regions when both the GOSAT and surface flask observations are used together. In a couple of cases though, the uncertainty reduction does not necessarily correspond with the GOSAT+SF flux estimates being nudged in the right direction. A typical example is Temperate Asia in January 2010, where a reduction in uncertainty (17%; Figure 7.5B) did not necessarily correspond with the flux estimate being any closer to the bottom-up estimate (TeAs; Figure 7.7B). Compared to a CO_2 source that is inferred from the SF-only inversion, the GOSAT+SF inversion inferred a small sink over this region (Figure 7.3B – parts of Xinjiang Province in China, Tibet and India). Given the limited biospheric activity taking place over this region during the winter season, a carbon sink is highly inconsistent with the underlying physical processes regulating the CO_2 fluxes. Biases in the GOSAT glint retrievals (e.g. *O'Dell et al*. [2012]) have been reported, and it is possible that these biases adjust the GOSAT+SF inversion in the wrong direction.

For both Figures 7.6 and 7.7, alongside the flux estimates from the GOSAT +SF and the SF-only inversion, flux estimates based on a GOSAT-only inversion are also shown. The GOSAT+SF inversion flux estimates lies in between the GOSAT-only and the SF-only inversion estimated fluxes. Over several regions, for example, the Tropical ocean regions in June 2009 or the Southern ocean regions in January 2010, the GOSAT+SF inversion flux estimates are more consistent with the bottom-up estimate than either the SF-only or the GOSAT-only flux estimates. If the GOSAT observations are biased, however, the estimated fluxes from the GOSAT+SF inversion degrade significantly. Over North Africa for both June 2009 and January 2010 (NoAf; Figure 7.7), GOSAT operates primarily in glint mode and quite likely biases in the

glint retrievals inform the inversion of a wrong source estimate over this region. If the GOSAT retrievals had large uncertainty associated with them, the inversion would have given less weight to these observations relative to the surface flask data. For example, over Boreal Asia (BoAs; Figure 7.7A), the surface flask sites are located over continental areas, and have a large model-data mismatch uncertainty (3.0 ppm) associated with them. As a consequence for the GOSAT+SF inversion, less weight is given to the imprecise flask observations while more information is extracted from the GOSAT observations. Unlike the flux estimates from the SF-only inversion, the estimated fluxes from the GOSAT+SF inversion are not impacted negatively by the surface flask data. Overall, the ability of GOSAT to provide strong constraints on the surface fluxes relative to the surface flask network is dependent on both the number of available retrievals over a region, and the quality of the retrievals. These different constraints provided by the GOSAT and the surface flask observations are assessed more quantitatively in the next section using the information matrix analysis.

7.4.3 Analysis sensitivity with respect to the observations

Figure 7.8 shows the analysis self-sensitivity ($\mathbf{S}^{o}_{ii}$) for each of the GOSAT and surface flask observations for the two examined months. Qualitatively, it is evident that the surface flask observations provide the largest source of information to the analysis for both months. Quite likely this is due to the fact that the GOSAT column-averaged observations are less sensitive to the surface fluxes than the surface flask observations. For the examined months, the influence of the satellite and the surface observations varies between 0.02-0.04 and 0.24-0.25, respectively. The relatively small influence of the surface flask and the GOSAT observations indicate that only a small portion of the information in the analysis comes from the observation at each observation location. Within the GEnSRF framework, the remaining information to the analysis

is provided by the ensemble. At the very first assimilation time step, the ensemble is representative of the information contained in the prior error covariance matrix. Over subsequent time steps, the ensemble accumulates information from the assimilated observations, and a portion of this information is used by the analysis. Note that within a Bayesian DA framework, however, the prior flux estimates (Bayesian framework) based on a bottom-up model and/or inventory, would have provided the background information to the analysis. But given the relatively low influence of the observations, any spatial patterns or assumptions associated with the bottom-up model would have significantly impacted the estimated fluxes. Thus, relative to existing Bayesian CO_2-DA systems, the value of GEnSRF lies in providing stronger data-driven estimates of surface fluxes, which can be used to assess the impact of the observational quality and quantity of the GOSAT retrievals in an independent fashion.

The influence of the observations increases if either the measurement error or the total model-data mismatch variances are reduced. For example, if all the surface flask observations are specified to have a low model-data mismatch uncertainty of 0.5 ppm (i.e., 0.25 ppm^2), then the average influence of the surface flask observations increases to 0.60-0.65. Additionally, the observational influence is also dependent on the observation density within a region. These can be seen by comparing the influence of the surface flask observations over the Southern Pacific in June (Figure 7.8A) and January (Figure 7.8B). Since the surface flask observations have the same model-data mismatch error for both the months, the reduction in their influence can be directly attributed to the increase in the number of GOSAT retrievals over this region in January. Likewise, the influence of the GOSAT observations is reduced in areas, where a large number of retrievals are available. For example, over Australia in June 2009 where the GOSAT retrievals are dense, the self-sensitivity of each observation tends to be small in order to reduce the relative

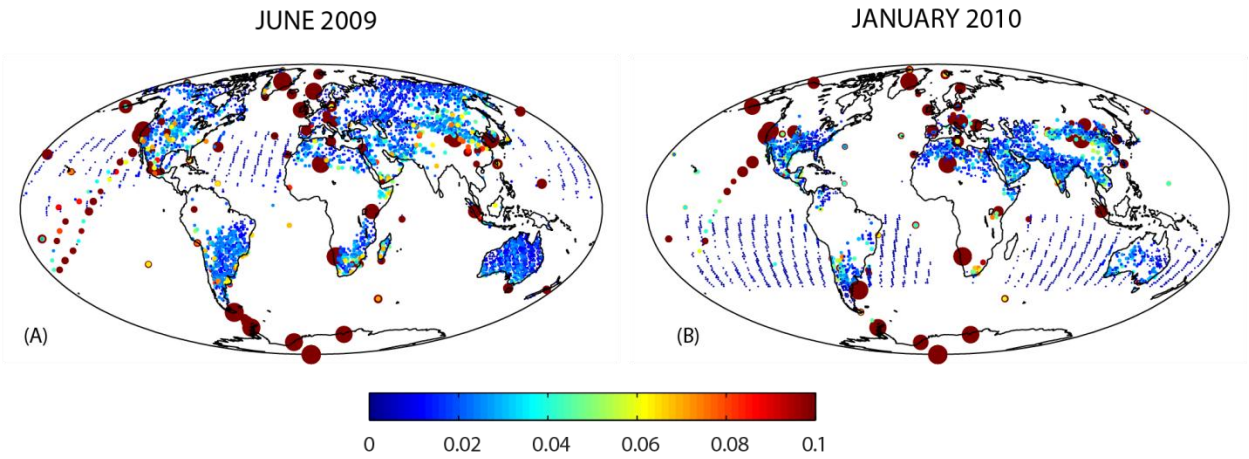


Figure 7.8– *Observational influence of GOSAT XCO_2 and surface flask CO_2 observations for June 2009 and January 2010. Higher the observational influence of a site, larger is the size of the circles plotted at that site. Typically the observational influence is plotted on a scale of 0 to 1 but here the color scale goes from 0 to 0.1 to bring out specifically the influence of the GOSAT XCO_2 observations.*

impact of the same type of observations on the analysis, whereas the surrounding observations from the surface flask become more influential.

Figure 7.8 demonstrates that, for estimation of surface fluxes, the average influence of the GOSAT data is significantly lower relative to the influence of the surface flask data. The GOSAT data are primarily able to contribute to the analysis because of their high density and good spatial coverage over large areas of the globe that are unobserved by the surface flask network. The limited sensitivity of the GOSAT observations to the surface fluxes can be either due to existing retrieval errors (e.g. *Takagi et al*. [2011]) or simply due to the inability of the transport model to correctly link the variations in the column-averaged CO_2 concentrations to the changes in the surface fluxes (e.g. *Houweling et al*. [2010]). The performance of atmospheric transport models and the impact of transport model errors on flux inversions using satellite measurements remains a high-priority research area within the larger CO_2 community as well (e.g. *Houweling et al*. [2010]; *Chevallier et al*. [2010b]).

7.5 SUMMARY

This study presented the results of assimilating GOSAT XCO_2 observations into a geostatistical ensemble square root filter (GEnSRF) to obtain high resolution (spatial - 1° × 1.25° and temporal – daily) estimates of CO_2 surface fluxes. Alongside the GOSAT observations, atmospheric CO_2 observations from an *in situ* monitoring network are also assimilated. Based on the experiments conducted in this study, it is evident that the GOSAT and the surface flask observations have complementary benefits. While the surface flask observations are more sensitive to the surface fluxes, the GOSAT data provides better coverage in space and time. Typical regions that are unconstrained by the surface network (e.g. the Tropical and the Southern Ocean regions), are well observed by GOSAT. Being able to obtain flux estimates over these regions fills a substantial gap in our knowledge of regional and global carbon budgets. Hence, using both the GOSAT and the surface flask observations in synergy is a better option for retrieving CO_2 surface fluxes than using observations from a single instrument. Quantification of the observational influence indicates that the high sensitivity of the ground-based network to the surface fluxes play a greater role in influencing the analysis than the satellite observations. As the GOSAT XCO_2 retrievals mature and the data product becomes more precise, the influence of the GOSAT observations will likely increase. Consequently, the recovered CO_2 fluxes should aid in better qualitative and quantitative definition of the processes driving the net uptake and release of carbon, especially over regions that were previously unconstrained.

The GEnSRF framework implemented in this study takes into account the versatility of the geostatistical inverse modeling as well as the computational efficiency of DA. GEnSRF provides the framework required to assimilate large volumes of satellite observation of CO_2 and provides estimates that minimize the influence of assumptions inherent to bottom-up estimates.

The developed framework is beneficial for future global studies with additional satellite data, as well as for conducting Observing System Simulation Experiments (OSSEs) for upcoming sensors such as the OCO-2 (e.g. *Eldering et al*. [2012]) and the ASCENDS (e.g. *NRC* [2007]) mission. Furthermore the influence matrix analysis can be easily embedded into the GEnSRF framework to demonstrate the impact of the observations. This likely will be valuable in the assessment and design of these future sensors.

Ongoing work aims to examine the drivers of the inter-annual variability in the carbon budget by estimating fluxes for an entire year (i.e., June 2009 – May 2010). This will also allow a direct comparison with other groups involved in inferring fluxes using Bayesian CO_2-DA systems. The GEnSRF analysis will provide a valuable basis for comparison to these estimates, and may help explain the influence of prior model assumptions on recovered fluxes. Second, even though the GEnSRF DA system was able to glean source-sink estimates from the GOSAT data at high spatiotemporal scales, robust carbon budgets could not be obtained. Our current inability to capture small-scale features in the flux patterns could potentially be overcome by incorporating auxiliary environmental variables into the model of the trend. Work is underway to assess the ability of the column-averaged XCO_2 observations to identify relevant environmental variables as well as remodel the DA setup to ingest these auxiliary variables.

Finally, other setup choices such as the use of a particular atmospheric transport model will be potentially considered in the future. The GEnSRF DA framework has been set up in a flexible manner (Appendix D) such that the atmospheric transport model component can be swapped in and out easily. This will provide insights into the role of transport model errors on the flux estimates and examine potential transport model improvements that are necessary to make optimal use of satellite measurements for estimating CO_2 sources and sinks.

Chapter 8

Conclusions

As part of the dissertation, four specific components were examined, each of which contributes in its own way to obtaining reliable estimates of atmospheric CO_2 concentrations and/or fluxes using data assimilation. This concluding chapter summarizes the findings of each of the four major components of the dissertation, the overall contribution of the dissertation, and also suggests directions for future work.

8.1 CONTRIBUTION OF EACH DISSERTATION COMPONENT

8.1.1 Background error statistics for atmospheric CO_2 data assimilation

The first dissertation component developed a model for the background error statistics for use in the ECMWF 4D-VAR system for estimating global 4D fields of atmospheric CO_2 concentrations. A new approach was developed that used the difference between modeled CO_2 distributions (ΔCO_2) as a surrogate for the 'true' background error. The ΔCO_2 field enabled specification of representative background error statistics for atmospheric CO_2 data assimilation, which was shown to - 1) vary regionally and seasonally to better capture the changing degree of variability in the background CO_2 field, and 2) have a discernible impact on the analysis estimates by allowing observations to adjust predictions over a larger area. The specific contribution of this component has been towards providing a method for realistic estimation of the background error statistics. This study illustrates the improvement that can be made and

therefore makes the clear case for including a good estimate of surface flux errors and error correlations in any method that is used for estimating the background errors of an atmospheric CO_2 data assimilation system. On a broader level though, these results establish the necessity to adapt existing data assimilation methods and/or the components of a DA system for carbon science applications.

8.1.2 Inter-comparison of ensemble and variational data assimilation in the context of a CO_2 flux estimation problem

The second component carried out a comparative assessment of two advanced data assimilation approaches with a batch inverse modeling scheme, specifically for the atmospheric CO_2 inversion problem. The primary outcome was an evaluation of the DA methods keeping in mind the nature of the CO_2 inversion problem. Results demonstrated that the performance of the DA approaches depends on a complex interplay between the underlying numerical approximations, the lack of a dynamical model, and the information available from the observations. The 4D-VAR scheme is found to be slightly more robust for obtaining CO_2 flux estimates while the EnSRF scheme provided useful estimates of analysis error, which were not directly available from the 4D-VAR system. The framework outlined in this study may also be used in the future to pursue more advanced intercomparison efforts for a real CO_2 flux estimation problem. Prior to this study, efforts were undertaken to interact with diverse groups involved in carbon data assimilation research. The results of this study are expected to benefit the CO_2 data assimilation community as a whole, in recognizing the pros and cons of different methods. As such, the conclusions and the framework are also relevant to a broader spectrum of DA applications, where dynamical models for propagating the state vector may not be readily available.

8.1.3 Towards reliable ensemble Kalman filter estimates of CO_2 fluxes

The third component developed the mathematical framework for a geostatistical ensemble square root filter (GEnSRF) that leverages the information content of atmospheric CO_2 observations more strongly relative to existing CO_2-DA approaches. The primary goal in developing GEnSRF was: a) to provide a basis for determining carbon sources or sinks with acceptable accuracy at fine spatial and temporal scales, and b) to aid us in understanding the crucial controls on the global carbon cycle and its responses to anthropogenic forcing. Within this context, the ability of the GEnSRF to yield reliable flux estimates and uncertainties across a range of resolutions over North America, without the use of prior fluxes from any bottom-up estimate, was found to be extremely promising. Several advanced techniques, such as adaptive covariance inflation and localization were implemented to minimize well-known error sources in the ensemble filter, and the sensitivity of the filter to these algorithms tested as well. Even though the primary application in this study was in an atmospheric CO_2 inverse modeling context, GEnSRF remains applicable to complementary inverse problems, for example constraining sources and sinks of other trace gases of interest like methane, CFCs and N_2O.

8.1.4 Role of GOSAT total column CO_2 observations for the estimation of CO_2 surface fluxes

The final component used GEnSRF to estimate global carbon fluxes at a high spatial and temporal resolution, using solely the atmospheric data constraint provided by a combination of GOSAT XCO_2 retrievals and atmospheric CO_2 observations from a conventional surface flask network. A state-of-the-art data assimilation infrastructure was developed that provided not only knowledge about the global fluxes but also evaluated the role of GOSAT observations in carbon cycle science. The data-driven estimates are not influenced by *a priori* assumptions from bottom-

up flux estimates, and thus directly reflect the information content of the GOSAT observations. Even though the satellite observations are found to be useful in constraining ocean fluxes, where the surface network has limited coverage, they have a limited influence on the analysis relative to the ground-based network. This can be attributed to both the limited sensitivity of the satellite observations to the surface fluxes as well as the inability of current atmospheric transport models to link variations in column-averaged CO_2 concentrations to the variations in surface fluxes. Thus, going forward a high priority research area for the CO_2-DA community should be in evaluating ways to possibly extract more information from the high-density GOSAT data than currently possible. Additionally, since GEnSRF provides a computationally efficient platform for examining the analysis sensitivity with respect to the observations, it has a lot of potential for the design and assessment of future CO_2 observing systems.

8.2 OVERALL CONTRIBUTION OF DISSERTATION

The primary contribution of this dissertation has been in identifying fundamental gaps in our understanding/applications of CO_2-DA, and developing suitable methods to fill these gaps. The sum of the different components of the dissertation increases the credibility of both existing and new CO_2-DA systems. Several of the key issues examined in this dissertation (i.e. impact of background error statistics, lack of a dynamical flux model, impact of numerical approximations, impact of observational network constraints, etc.) have been primarily motivated by the uncertainty surrounding the applicability and accuracy of DA approaches for atmospheric CO_2 applications. The attempt throughout has been to reduce this uncertainty and develop improved CO_2-DA tools, which may ultimately aid in understanding the critical controls over the atmospheric CO_2 growth.

As the cost of carbon observing systems is rising, the need for efficiently using the benefits of new observations, such as expanded CO_2 surface networks, and satellite observations, is steadily increasing. Simultaneously, rapid advancements are being made in the development of process-based models of the carbon cycle. The development of GEnSRF is especially beneficial in this regard, as it minimizes process-based assumptions inherent to other existing CO_2-DA systems and instead allows validation of the process-based models by independent data-driven estimates of carbon flux. In fact, by determining as accurately as possible the state of the CO_2 fluxes, the geostatistical data assimilation system may also be used in the future for the monitoring and verification of fossil and bio fuels emissions and sequestration.

Alongside the scientific-cum-mathematical contribution, the DA infrastructure that has been set up as part of Chapter 7, in terms of the data processing, the coupling between the ensemble filter and the transport, and the general ensemble filter framework, is highly beneficial in its own right. It retains the ability to integrate multiple atmospheric, oceanic and terrestrial data sources seamlessly as may be necessary in the future with launch of new dedicated missions such as OCO-2 (*Eldering et al.* [2012]) and ASCENDS (*NRC* [2007]). Over the next decade, the greater coverage in time and space provided by the satellite data will yield new constraints on CO_2 fluxes that may dramatically improve our understanding of the surface CO_2 exchange. The infrastructure developed as part of this dissertation is the first step towards utilizing the full potential of a geostatistical DA system in efficiently solving the atmospheric CO_2 inverse problem.

8.3 FUTURE RESEARCH DIRECTIONS

This section describes avenues for future research based on the finding of this dissertation. The framework for this discussion derives from the perceived drawbacks and/or

unexplored areas in CO_2-DA applications. In reality the range of possible directions for the data assimilation work is indescribably varied due to the large number of DA tools that can be used, and the subtle methodological advancements that can be made within each of these tools. This section will focus mainly on ensemble filter applications, which are arguably becoming more popular than other DA approaches due to their ease of operational implementation. Unfortunately, this also implies that a discussion on 4D-VAR and/or hybrid schemes remain beyond the scope of this chapter.

8.3.1 Direct extensions to current work

While the GEnSRF tool developed for this dissertation is starting to yield promising flux estimates at global scales, an important issue that has not been addressed is the impact of auxiliary variables on the flux estimates. Recall that the geostatistical framework (Chapter 3 – Section 3.3.2) allows for inclusion of auxiliary environmental data related to CO_2 flux to help constrain the estimates. These additional variables help to capture the small scale variability in the flux field (e.g. *Gourdji et al*. [2008]) than what is visible through the atmospheric data. An added advantage of including these covariates is that, if the model of the trend (i.e., $\mathbf{X\beta}$) is able to explain more of the variability in the recovered fluxes, then the influence of the spatiotemporal correlation structure in $\mathbf{Q}^b$ is significantly reduced in the final estimates. Inclusion of auxiliary variables in the GEnSRF DA may require modifications of the existing mathematical framework but will definitely improve the ability of the GEnSRF to provide more realistic flux estimates and uncertainties.

An associated issue is the fact that there is no dynamical (or time evolution) model to evolve the CO_2 flux (or state vector) forward in time. The lack of a dynamical model is clearly one of the main challenges specific to the CO_2 flux estimation problem but is most evident,

however, when operational constraints are imposed (*Chatterjee and Michalak* [in prep.]). Using the information from the auxiliary variables to substitute for the missing dynamical model may inject valuable physical information into the DA system. It is important to remember that GEnSRF has a data assimilation system at its core. Thus, if a suitable substitute for the dynamical model can be identified and/or specified, the existing GEnSRF setup can be easily modified to work as a prediction tool.

Secondly, specific to all ensemble filter applications for CO_2 an important issue is the impact of the covariance localization scheme and the specified localization length scales. Currently, all ensemble filter applications for CO_2 specify a compactly supported function to filter out the sampling error noise. Studies (e.g. *Lokupitiya et al*. [2008]; *Chatterjee et al*. [in press]) have pointed out that the atmospheric advection of CO_2 may not be consistent with the use of compactly supported correlation functions. Apart from the compactly supported functions, an alternate option for CO_2 applications is to use the dynamical localization scheme as suggested by *Zupanski et al*. [2007a]. In this approach, the distance for covariance localization is not based on the geodesic distance but rather the ratio between the prior and the posterior uncertainty. A significant challenge though is coming up with an estimate of the posterior uncertainty prior to the analysis. The above study is based on a hybrid Maximum Likelihood Ensemble Framework (MLEF) framework, within which the posterior uncertainty is available as a byproduct of the Hessian pre-conditioning (*Zupanski* [2005]). The 4D-VAR implementation allows this but it is less clear how to calculate this within a regular ensemble framework.

In addition, choosing the localization length scale remains rather subjective (e.g. *Chatterjee et al*. [in press]) and is based on a number of sensitivity tests. This is complicated by the fact that the localization length scale itself depends on the ensemble size and the number of

observations used in the assimilation. Recently *Anderson* [2012] has proposed a sampling error correction technique in which the localization length scale is calculated based on the prior ensemble. This method does not use observational information, however, and will not be able to correct for systematic biases in the model or other forms of model error. Localization is one of key aspects within the ensemble filter and identifying a suitable approach for obtaining realistic localization length scales will significantly improve ensemble filter based CO_2 estimates, including those from GEnSRF.

Thirdly, atmospheric CO_2 observations are not sufficient to understand all the different aspects of the global carbon cycle. In reality, atmospheric CO_2 observations can provide reliable constraints on the net flux budget but not necessarily on the key signatures of anthropogenic and/or biogenic processes driving those fluxes. A possible way of inferring information about the processes that influence the surface fluxes of CO_2 may be by examining or incorporating atmospheric measurements of process-specific tracers in the inversion process. To this end, there may be value in assimilating heterogeneous observations of carbon tracer species within an integrated earth system model using a data assimilation framework. Development and implementation of this multi-species carbon data assimilation framework will be the basis for my future research at the National Centre for Atmospheric Research, as part of a NOAA Climate and Global Change Postdoctoral Fellowship (Appendix E). The aim of this research will be to merge information of atmospheric trace gas species of carbon with CO_2 measurements with state of the art community models using an ensemble smoother. Overall, this study will explore the value of adopting a multispecies approach for improved CO_2 source attribution including fingerprinting the role of anthropogenic forcing on carbon cycle dynamics.

8.3.2 Larger community-wide directions for improving CO_2-DA applications

Before delving into data assimilation related improvements, it is worthwhile to briefly mention one aspect, which is more relevant to the CO_2 inverse modeling community as a whole. There is an immediate need to improve transport modeling capabilities in order to: (a) eliminate systematic biases that may affect inversion results, (b) allow better use of night-time CO_2 measurements (see Chapter 6) or measurements retrieved over complex terrains (e.g. *van der Molen and Dolman* [2007]; *Brooks et al*. [2012]) from continuous sites, and (c) make better use of satellite data by correctly simulating key processes such as the time scales of stratosphere-troposphere exchange and planetary boundary layer dynamics (e.g. *Houweling et al*. [2010]; *Chevallier et al*. [2010b]). Recent efforts have been made in improving certain aspects of the transport model, for example the representation of vertical transport within the planetary boundary layer (e.g. *Gerbig et al*. [2008]; *Kretschmer et al*. [2012]) but the community still lacks a pragmatic knowledge of the real magnitude of the transport model uncertainties. Clearly large-scale efforts (i.e., sustained observations, funding and improved theory) are necessary to improve this critical piece within an atmospheric CO_2 inverse modeling framework.

Finally, within the next decade I believe the CO_2-DA community will have to address two critical issues, if there is any hope in advancing DA for carbon science and/or utilizing the true capabilities of DA in understanding carbon-climate feedback issues.

First, it will be prudent for the CO_2-DA community to transition from an *inversion* to an *assimilation* framework. The inversion framework loses out on the predictive capabilities associated with a true DA framework as well as the ability to directly adjust the parameters of biogeochemical model components using atmospheric CO_2 observations (*Rayner* [2010]). In an assimilation framework it is quite feasible to identify which parameters of the biogeochemical

model are less well-understood, and subsequently design an observational network/system to constrain those parameters. A research framework adopting an integrated observation and modeling approach may likely provide the scientific basis for future atmospheric CO_2 mitigation strategies, but it will require a community wide effort to advance current CO_2-DA flux estimation systems to CO_2-DA prediction systems.

A second issue will be to generate more interest and expertise within the CO_2 community in developing and implementing DA tools. Surprisingly, this issue is not unique to the CO_2 field and has been identified within the larger DA community as well (*Vukicevic et al*. [2004]). Over the past decade, the NWP community has made concerted efforts to organize educational and outreach DA workshops to train young scientists and professionals, in not just working with existing DA systems but interested in development of assimilation tools. The CO_2-DA community may as well follow the NWP paradigm (e.g. *Roebber et al*. [2010]) in providing more exposure to young scientists and graduate students with systems of relevant complexity and with relevant computational infrastructure. This will require the CO_2-DA community, however, to first graduate from working on independent research agendas and independent DA systems to developing a more multi-faceted DA system with strong operational components.

Appendix A

Table A1*- Reference list of mathematical symbols used in this dissertation.*

Symbols	Name	Dimensions
$\mathbf{s}$	state vector	$m \times 1$
$\mathbf{s}'$	state vector deviations	$m \times N$
$\mathbf{s}^b$	prior (or background)estimate of the state vector	$m \times 1$
$\mathbf{Q}^b$	prior (or background) error covariance matrix	$m \times m$
$\hat{\mathbf{s}}$	posterior best estimate of $\mathbf{s}$	$m \times 1$
$\mathbf{Q}^a$	*a posteriori* covariance of $\mathbf{s}$	$m \times m$
$\mathbf{X}$	drift (or trend) from auxiliary variables	$m \times p$
$\boldsymbol{\beta}$	drift coefficients defining the weight to each variable	$p \times 1$
$\mathbf{z}$	observation vector	$n \times 1$
$\mathbf{R}$	observation error covariance matrix	$n \times n$
h	forward model operator	
$\mathbf{H}$	matrix form of h (also called sensitivity matrix)	$n \times m$
M	forecast/dynamical model operator	
$\mathbf{K}$	Kalman gain	$m \times n$
$\tilde{\mathbf{K}}$	reduced Kalman gain	$m \times n$
$\boldsymbol{\Lambda}$	matrix of coefficients	$m \times n$
$\tilde{\boldsymbol{\Lambda}}$	reduced matrix of coefficients	$m \times n$
$\mathbf{M}$	matrix of Lagrange multipliers	$p \times m$
$\mathbf{d}$	separation distance between grid points in space	
t	time	
γ	semi-variance at separation distance h	
σ^2	sill parameter of a variogram	
l	range parameter of a variogram	
i, j, k, τ	generic integer indices	
$J(\cdot)$	objective function (also called cost or penalty function)	
∇	Gradient operator	

Symbols	Name	Dimensions
$(\cdot)^f$	Forecast values	
$(\cdot)^a$	Analyzed values	
$(\cdot)^b$	Background values	
$(\cdot)^T$	Transpose operator	
$(\cdot)^{-1}$	Inverse operator	

Appendix B

***Table B1**- Aircraft (blue) and TCCON (purple) sites used in the evaluation of the 4D-VAR analysis in Chapter 4. Data availability at a site is denoted with a 'Y'. Note that all the aircraft sites are located over North America.*

Measurement site code	Site name	Site latitude/ longitude (degrees)	Data for January 2010	Data for June 2010
BNE	Beaver Crossing, Nebraska	40.80 N, 97.18 W	Y	Y
CAR	Briggsdale, Colorado	40.37 N, 104.30 W	Y	Y
CMA	Cape May, New Jersey	38.83 N, 74.32 W	Y	Y
DND	Dahlen, North Dakota	48.38 N, 99.00 W	-	Y
ESP	Estevan Point, B. Columbia	49.58 N, 126.37 W	Y	Y
ETL	East trout Lake, Saskatchewan	54.35 N, 104.98 W	Y	Y
HIL	Homer, Illinois	40.07 N, 87.91 W	Y	Y
LEF	Park Falls, Wisconsin	45.95 N, 90.27 W	Y	Y
NHA	Worcester, Massachusetts	42.95 N, 70.63 W	Y	Y
PFA	Poker Flats, Alaska	65.07 N, 147.29 W	Y	Y
SCA	Charleston, South Carolina	32.77 N, 79.55 W	Y	Y
SGP	Southern Great Plains, Oklahoma	36.80 N, 97.50 W	Y	Y
TGC	Sinton, Texas	27.73 N, 96.86 W	Y	Y
THD	Trinidad Head, California	41.05 N, 124. 15 W	Y	Y
WBI	West Branch, Iowa	41.72 N, 91.35 W	Y	Y
BIA	Bialystok, Poland	53.23 N, 23.03 E	Y	Y
BRE	Bremen, Germany	53.10 N, 8.85 E	Y	Y
DAR	Darwin, Australia	12.42 S, 130.89 E	Y	-
EUR	Eureka, Canada	80.05 N, 86.42 W	-	-
GAR	Garmisch, Germany	47.48 N, 11.06 E	Y	Y
IZA	Izana, Tenerife	28.30 N, 16.50 W	-	Y
KAR	Karlsruhe, Germany	49.10 N, 8.44 E	-	Y
LAU	Lauder, New Zealand	45.04 S, 169.68 E	-	Y

Measurement site code	Site name	Site latitude/ longitude (degrees)	Data for January 2010	Data for June 2010
LEF	Park Falls, Wisconsin	45.95 N, 90.27 W	Y	-
NYA	NY Alesund, Norway	78.92 N, 11.92 E	-	Y
ORL	Orleans, France	47.97 N, 2.11 E	-	Y
SGP	Lamont, Oklahoma	36.60 N, 97.49 W	Y	Y
SOD	Sodankyla, Finland	67.37 N, 26.63 E	-	Y
TSU	Tsukuba, japan	36.05 N, 140.12 E	Y	-
WOL	Wollongong, Australia	34.41 S, 150.88 E	Y	-

Appendix C

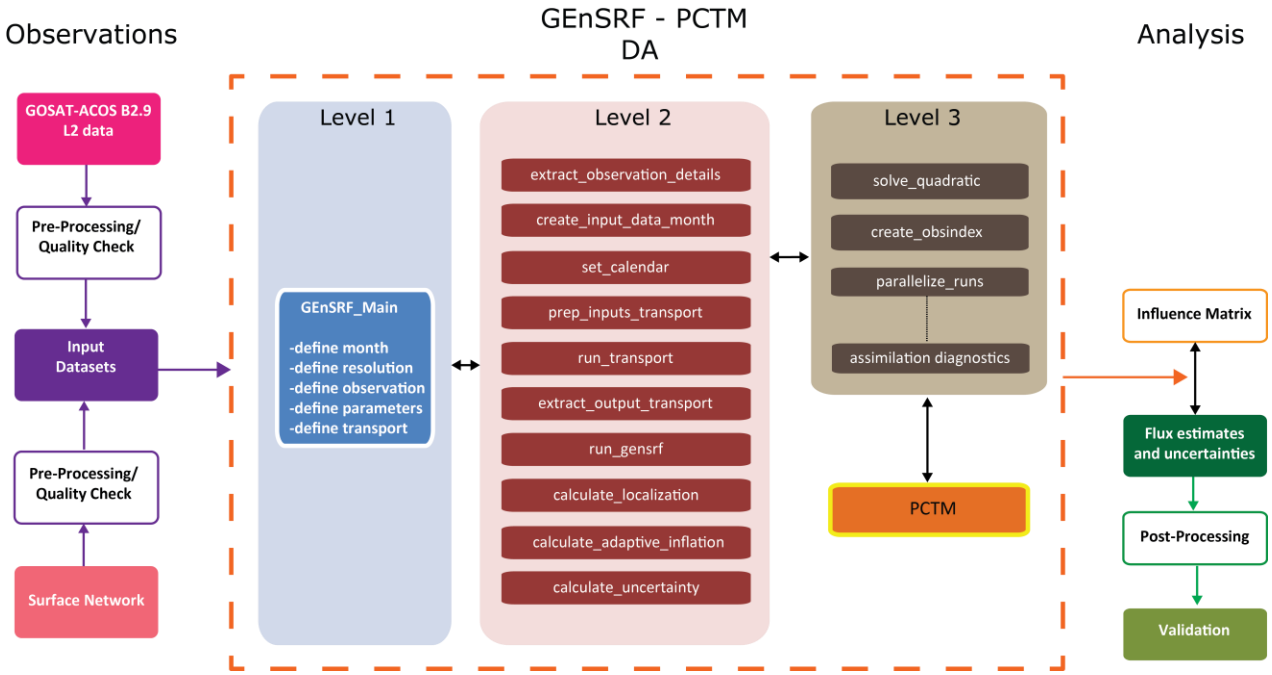
***Table C1**- Measurement network used in the inversions in Chapter 6.*

Measurement site code	Site name	Site latitude/ longitude (degrees)	Height above ground level (m)
LEF	Park Falls, Wisconsin	45.93N, 90.27W	396
WKT	Moody, Texas	31.32N, 97.33W	457
WBI	West Branch, Iowa	41.73 N, 91.35 W	379
BAO	Boulder Atmospheric Obs., Colorado	40.05 N, 105.01 W	300
SCT	South Carolina Tower, South Carolina	33.41 N, 81.83 W	305
WGC	Walnut Grove, California	38.27 N, 121.49 W	483
AMT	Argyle, Maine	45.03N, 68.68W	107
BRW	Barrow, Alaska	71.32N, 156.60W	10
FRD	Fraserdale, Ontario	49.88 N, 81.57 W	40
CDL	Candle Lake, Saskatchewan	53.99N, 105.12W	30
SBL	Sable Island, Nova Scotia	43.93N, 60.02W	25
EGB	Egbert, Ontario	44.23 N, 79.78 W	3
ETL	Saskatchewan, East Trout Lake	54.35 N, 104.99 W	105
LLB	Lac LaBich, Alberta	54.95 N, 112.45 W	10
CHI	Chibougamau, Quebec	49.69 N, 74.34 W	30
HFM	Harvard Forest, Massachusetts	42.54N, 72.17W	30
ARM	Norman, Oklahoma	36.80 N, 97.50W	60
CVA	Canaan Valley, West Virginia	39.06 N, 79.42 W	7
MOM	Morgan Monroe, Indiana	39.32 N, 86.41 W	48
OZA	Ozark, Missouri	38.74 N, 92.20 W	30
KEW	Kewanee, Illinois	41.28 N, 89.97 N	140
CEN	Centerville, Iowa	40.79 N, 92.88 W	110
MEA	Mead, Nebraska	41.14 N, 96.46 W	122

Measurement site code	Site name	Site latitude/ longitude	Height above ground level (m)
ROL	Round Lake, Minnesota	43.53 N, 95.41 W	110
GAL	Galesville, Wisconsin	44.09 N, 91.34 W	122
SNP	Shenandoah National Park, Virginia	38.62 N, 78.35 W	17
SPL	Storm Peak Lab, Colorado	40.45 N, 106.73 W	9
NWR	Niwot Ridge, Colorado	40.05 N, 105.58 W	5
HDP	Hidden Peak Snowbird, Utah	40.56 N, 111.65 W	18
FIR	Fir, Oregon	44.65 N, 123.55 W	38
MET	Metolius, Oregon	44.45 N, 121.56 W	34
YAH	Yaquina Head, Oregon	44.67 N, 124.07 W	13
MAP	Mary's Peak, Oregon	44.50 N, 123.55 W	8
NGB	NGBER, Oregon	43.47 N, 119.69 W	7
LJA	LaJolla, California	32.87 N, 117.26 W	5

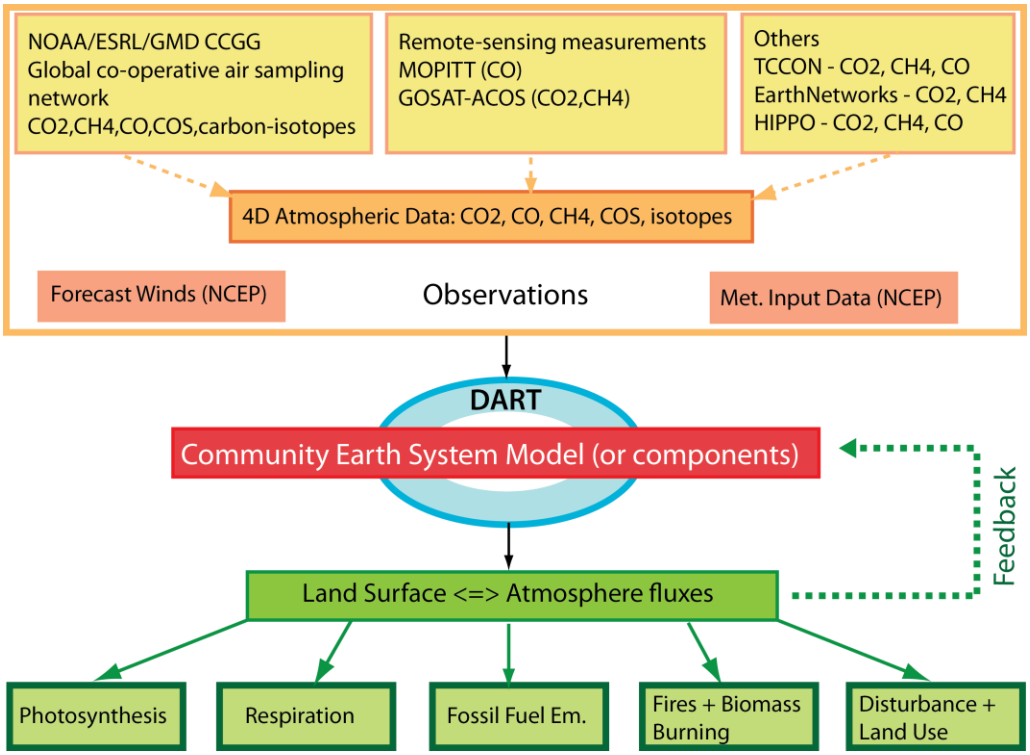
Appendix D

Figure D1- *Flow of the data assimilation system used in Chapter 7, showing the interactions between different ingredients of the problem such as the observations, the ensemble filter, the transport scheme etc. Based on a shared-memory multiprocessing framework, the system uses a geostatistical ensemble square root filter to assimilate observations from different sources and generate estimates of global land and oceanic CO_2 exchange at fine spatiotemporal scales. Ultimately, this system will be deployed as part of a larger parallel autonomous software platform (Yadav et al. [2010b]) for real-time integration of in-situ and satellite-based atmospheric CO_2 observations.*



Appendix E

Figure E1*- Schematic of a proposed data assimilation framework to investigate:* 1) *the value in assimilating information from disparate carbon tracers for reliable source attribution, and 2) potential improvements to carbon cycle models and their predictive capabilities by assimilation of a variety of carbon tracers. Observations of carbon tracers from different platforms will be merged with the components of the Community Earth System Model (CESM; http://www.cesm.ucar.edu/) using a n ensemble smoother that will be developed as part of a community data assimilation facility (DART; http://www.image.ucar.edu/DAReS/DART/#). Note that for brevity, sample observational databases for the different tracers have been highlighted. Appropriate datasets will be identified during the course of this work in discussion with different observational groups at the National Centre for Atmospheric Research (NCAR), and elsewhere.*



REFERENCES

Akaike, H. (1974), New Look at Statistical-Model Identification, *Ieee Transactions on Automatic Control*, *AC19*(6), doi:10.1109/tac.1974.1100705.

Alfieri, J. G., W. P. Kustas, J. H. Prueger, L. E. Hipps, J. L. Chavez, A. N. French, and S. R. Evett (2011), Intercomparison of Nine Micrometeorological Stations during the BEAREX08 Field Campaign, *Journal of Atmospheric and Oceanic Technology*, *28*(11), doi:10.1175/2011jtech1514.1.

Alkhaled, A. A., A. M. Michalak, S. R. Kawa, S. C. Olsen, and J.-W. Wang (2008), A global evaluation of the regional spatial variability of column integrated CO(2) distributions, *Journal of Geophysical Research-Atmospheres*, *113*(D20), doi:10.1029/2007jd009693.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano (2009), The Data Assimilation Research Testbed, A Community Facility, *Bulletin of the American Meteorological Society*, *90*(9), 1283-1296, doi:10.1175/2009bams2618.1.

Anderson, J. L. (2001), An ensemble adjustment Kalman filter for data assimilation, *Monthly Weather Review*, *129*(12), 2884-2903, doi:10.1175/1520-0493(2001)129<2884:aeakff>2.0.co;2.

Anderson, J. L. (2003), A local least squares framework for ensemble filtering, *Monthly Weather Review*, *131*(4), 634-642, doi:10.1175/1520-0493(2003)131<0634:allsff>2.0.co;2.

Anderson, J. L. (2007a), An adaptive covariance inflation error correction algorithm for ensemble filters, *Tellus Series a-Dynamic Meteorology and Oceanography*, *59*(2), 210-224, doi:10.1111/j.1600-0870.2006.00216.x.

Anderson, J. L. (2007b), Exploring the need for localization in ensemble data assimilation using a hierarchical ensemble filter, *Physica D-Nonlinear Phenomena*, *230*(1-2), doi:10.1016/j.physd.2006.02.011.

Anderson, J. L. (2009), Spatially and temporally varying adaptive covariance inflation for ensemble filters, *Tellus Series a-Dynamic Meteorology and Oceanography*, *61*(1), 72-83, doi:10.1111/j.1600-0870.2008.00361.x.

Anderson, J. L. (2012), Localization and Sampling Error Correction in Ensemble Kalman Filter Data Assimilation, *Monthly Weather Review*, *140*(7), doi:10.1175/mwr-d-11-00013.1.

Anderson, J. L., and S. L. Anderson (1999), A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, *Monthly Weather Review*, *127*(12), 2741-2758, doi:10.1175/1520-0493(1999)127<2741:amciot>2.0.co;2.

Andres, R. J., et al. (2012), A synthesis of carbon dioxide emissions from fossil-fuel combustion, *Biogeosciences*, *9*(5), 1845-1871, doi:10.5194/bg-9-1845-2012.

Arrhenius, S. (1896), On the Influence of Carbonic Acid in the Air Upon the Temperature of the Ground, *Philosophical Magazine*, 41, 237-76.

Auroux, D. (2007), Generalization of the dual variational data assimilation algorithm to a nonlinear layered quasi-geostrophic ocean model, *Inverse Problems*, *23*(6), doi:10.1088/0266-5611/23/6/013.

Baker, D. F., H. Boesch, S. C. Doney, D. O'Brien, and D. S. Schimel (2010), Carbon source/sink information provided by column CO(2) measurements from the Orbiting Carbon Observatory, *Atmospheric Chemistry and Physics*, *10*(9), 4145-4165, doi:10.5194/acp-10-4145-2010.

Baker, D. F., S. C. Doney, and D. S. Schimel (2006a), Variational data assimilation for atmospheric CO(2), *Tellus Series B-Chemical and Physical Meteorology*, *58*(5), 359-365, doi:10.1111/j.1600-0889.2006.00218.x.

Baker, D. F., et al. (2006b), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO2 fluxes, 1988-2003, *Global Biogeochemical Cycles*, *20*(1), doi:10.1029/2004gb002439.

Baker, I. T., L. Prihodko, A. S. Denning, M. Goulden, S. Miller, and H. R. da Rocha (2008), Seasonal drought stress in the Amazon: Reconciling models and observations, *Journal of Geophysical Research-Biogeosciences*, *113*, 10, doi:10.1029/2007jg000644.

Bakwin, P. S., P. P. Tans, D. F. Hurst and C. L. Zhao, (1998), Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, *Tellus B*, 50B, 5, 401-415, doi: 10.1034/j.1600-0889.1998.t01-4-00001.x

Baldocchi, D. (2008), Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems, *Australian Journal of Botany*, *56*(1), 1-26, doi:10.1071/bt07151.

Baldocchi, D., et al. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bulletin of the American Meteorological Society*, *82*(11), doi:10.1175/1520-0477(2001)082<2415:fantts>2.3.co;2.

Baldocchi, D. D. (2003), Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, *Global Change Biology*, *9*(4), doi:10.1046/j.1365-2486.2003.00629.x.

Bannister, R. N. (2008a), A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, *Quarterly Journal of the Royal Meteorological Society*, *134*(637), 1951-1970, doi:10.1002/qj.339.

Bannister, R. N. (2008b), A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics, *Quarterly Journal of the Royal Meteorological Society*, *134*(637), 1971-1996, doi:10.1002/qj.340.

Bauer, P., P. Lopez, A. Benedetti, D. Salmond, and E. Moreau (2006), Implementation of 1D+4D-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1 D-Var, *Quarterly Journal of the Royal Meteorological Society*, *132*(620), doi:10.1256/qj.05.189.

Bauer, P., R. Buizza, C. Cardinali, and J.-N. Thepaut (2011), Impact of singular-vector-based satellite data thinning on NWP, *Quarterly Journal of the Royal Meteorological Society*, *137*(655), doi:10.1002/qj.733.

Beer, C., et al. (2010), Terrestrial Gross Carbon Dioxide Uptake: Global Distribution and Covariation with Climate, *Science*, *329*(5993), 834-838, doi:10.1126/science.1184984.

Bellard, C., C. Bertelsmeier, P. Leadley, W. Thuiller, and F. Courchamp (2012), Impacts of climate change on the future of biodiversity, *Ecology Letters*, *15*(4), 365-377, doi:10.1111/j.1461-0248.2011.01736.x.

Belo Pereira, M., and L. Berre (2006), The use of an ensemble approach to study the background error covariances in a global NWP model, *Monthly Weather Review*, *134*(9), 2466-2489, doi:10.1175/mwr3189.1.

Benedetti, A., and M. Fisher (2007), Background error statistics for aerosols, *Quarterly Journal of the Royal Meteorological Society*, *133*(623), 391-405, doi:10.1002/qj.37.

Bengtsson, T., C. Snyder, and D. Nychka (2003), Toward a nonlinear ensemble filter for high-dimensional systems, *Journal of Geophysical Research-Atmospheres*, *108*(D24), 10, doi:10.1029/2002jd002900.

Bergemann, K., and S. Reich (2010), A localization technique for ensemble Kalman filters, *Quarterly Journal of the Royal Meteorological Society*, *136*(648), 701-707, doi:10.1002/qj.591.

Berre, L., and G. Desroziers (2010), Filtering of Background Error Variances and Correlations by Local Spatial Averaging: A Review, *Monthly Weather Review*, *138*(10), 3693-3720, doi:10.1175/2010mwr3111.1.

Berre, L., S. E. Stefanescu, and M. B. Pereira (2006), The representation of the analysis effect in three error simulation techniques, *Tellus Series a-Dynamic Meteorology and Oceanography*, *58*(2), 196-209, doi:10.1111/j.1600-0870.2006.00165.x.

Berthelot, M., P. Friedlingstein, P. Ciais, J. L. Dufresne, and P. Monfray (2005), How uncertainties in future climate change predictions translate into future terrestrial carbon fluxes, *Global Change Biology*, *11*(6), 959-970, doi:10.1111/j.1365-2486.2005.00957.x.

Billesbach, D. P. (2011), Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method, *Agricultural and Forest Meteorology*, *151*(3), 394-405, doi:10.1016/j.agrformet.2010.12.001.

Birdsey, R. (2004), Data gaps for monitoring forest carbon in the united states: An inventory perspective, *Environmental Management*, *33*, doi:10.1007/s00267-003-9113-6.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar (2001), Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects, *Monthly Weather Review*, *129*(3), 420-436, doi:10.1175/1520-0493(2001)129<0420:aswtet>2.0.co;2.

Bishop, C. H., and D. Hodyss (2011), Adaptive Ensemble Covariance Localization in Ensemble 4D-VAR State Estimation, *Monthly Weather Review*, *139*(4), 1241-1255, doi:10.1175/2010mwr3403.1.

Bocquet, M. (2005), Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: Theory, *Quarterly Journal of the Royal Meteorological Society*, *131*(610), 2191-2208, doi:10.1256/qj.04.67.

Boesch, H., D. Baker, B. Connor, D. Crisp, and C. Miller (2011), Global Characterization of CO2 Column Retrievals from Shortwave-Infrared Satellite Observations of the Orbiting Carbon Observatory-2 Mission, *Remote Sensing*, *3*(2), doi:10.3390/rs3020270.

Boesch, H., et al. (2006), Space-based near-infrared CO2 measurements: Testing the Orbiting Carbon Observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin, *Journal of Geophysical Research-Atmospheres*, *111*(D23), doi:10.1029/2006jd007080.

Boloni, G., and K. Horvath (2010), Diagnosis and tuning of background error statistics in a variational data assimilation system, *Idojaras, Quarterly Journal of the Hungarian Meteorological Society*, *114*(1-2), 1-19.

Bonavita, M., L. Raynaud, and L. Isaksen (2011), Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: some effects of ensemble size and day-to-day variability, *Quarterly Journal of the Royal Meteorological Society*, *137*(655), 423-434, doi:10.1002/qj.756.

Bondeau, A., et al. (2007), Modelling the role of agriculture for the 20th century global terrestrial carbon balance, *Global Change Biology*, *13*(3), doi:10.1111/j.1365-2486.2006.01305.x.

Bousquet, P., P. Peylin, P. Ciais, C. Le Quere, P. Friedlingstein, and P. P. Tans (2000), Regional changes in carbon dioxide fluxes of land and oceans since 1980, *Science*, *290*(5495), doi:10.1126/science.290.5495.1342.

Bouttier, F. and P. Courtier (1999) Data Assimilation Concepts and Methods, *ECMWF Metereological Training Course Lecture Series*, Available at http://www.ecmwf.int/newsevents/training/lecture_notes/pdf_files/ASSIM/Ass_cons.pdf

Brankart, J.-M., C. Ubelmann, C.-E. Testut, E. Cosme, P. Brasseur, and J. Verron (2009), Efficient Parameterization of the Observation Error Covariance Matrix for Square Root or Ensemble Kalman Filters: Application to Ocean Altimetry, *Monthly Weather Review*, *137*(6), doi:10.1175/2008mwr2693.1.

Braswell, B. H., W. J. Sacks, E. Linder, and D. S. Schimel (2005), Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations, *Global Change Biology*, *11*(2), doi:10.1111/j.1365-2486.2005.00897.x.

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. Available at http://cdiac.esd.ornl.gov/ndps/ndp058a.html.

Brooks, B. G. J., A. R. Desai, B. B. Stephens, D. R. Bowling, S. P. Burns, A. S. Watt, S. L. Heck, and C. Sweeney (2012), Assessing filtering of mountaintop CO2 mole fractions for application to inverse models of biosphere-atmosphere carbon exchange, *Atmospheric Chemistry and Physics*, *12*(4), 2099-

2115, doi:10.5194/acp-12-2099-2012.
Brousseau, P., L. Berre, F. Bouttier, and G. Desroziers (2012), Flow-dependent background-error covariances for a convective-scale data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *138*(663), 310-322, doi:10.1002/qj.920.
Bruhwiler, L. M. P., A. M. Michalak, W. Peters, D. F. Baker, and P. Tans (2005), An improved Kalman Smoother for atmospheric inversions, *Atmospheric Chemistry and Physics*, *5*, 2691-2702, doi:10.5194/acp-5-2691-2005
Bruhwiler, L. M. P., A. M. Michalak, and P. P. Tans (2011), Spatial and temporal resolution of carbon flux estimates for 1983-2002, *Biogeosciences*, *8*(5), doi:10.5194/bg-8-1309-2011.
Buchwitz, M., et al. (2005), Atmospheric methane and carbon dioxide from SCIAMACHY satellite data: initial comparison with chemistry and transport models, *Atmospheric Chemistry and Physics*, *5*, 941-962, doi:10.5194/acp-5-941-2005
Buchwitz, M., O. Schneising, J. P. Burrows, H. Bovensmann, M. Reuter, and J. Notholt (2007), First direct observation of the atmospheric CO2 year-to-year increase from space, *Atmospheric Chemistry and Physics*, *7*(16), 4249-4256, doi:10.5194/acp-7-4249-2007
Buehner, M., and M. Charron (2007), Spectral and spatial localization of background-error correlations for data assimilation, *Quarterly Journal of the Royal Meteorological Society*, *133*(624), 615-630, doi:10.1002/qj.50.
Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He (2010a), Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments, *Monthly Weather Review*, *138*(5), 1550-1566, doi:10.1175/2009mwr3157.1.
Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He (2010b), Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations, *Monthly Weather Review*, *138*(5), 1567-1586, doi:10.1175/2009mwr3158.1.
Buehner, M., and A. Mahidjiba (2010), Sensitivity of Global Ensemble Forecasts to the Initial Ensemble Mean and Perturbations: Comparison of EnKF, Singular Vector, and 4D-Var Approaches, *Monthly Weather Review*, *138*(10), 3886-3904, doi:10.1175/2010mwr3296.1.
Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), Analysis scheme in the ensemble Kalman filter, *Monthly Weather Review*, *126*(6), 1719-1724, doi:10.1175/1520-0493(1998)126<1719:asitek>2.0.co;2.
Callendar, G. S. (1938), The artificial production of carbon dioxide and its influence on temperature, *Quarterly Journal of the Royal Meteorological Society*, *64* (275), 223-240, doi: 10.1002/qj.49706427503
Canadell, J. G., et al. (2000), Carbon metabolism of the terrestrial biosphere: A multitechnique approach for improved understanding, *Ecosystems*, *3*(2), 115-130, doi:10.1007/s100210000014.
Canadell, J. G., P. Ciais, P. Cox, and M. Heimann (2004), Quantifying, understanding and managing the carbon cycle in the next decades, *Climatic Change*, *67*(2-3), 147-160, doi:10.1007/s10584-004-3765-y.
Canadell, J. G., C. Le Quere, M. R. Raupach, C. B. Field, E. T. Buitenhuis, P. Ciais, T. J. Conway, N. P. Gillett, R. A. Houghton, and G. Marland (2007), Contributions to accelerating atmospheric CO(2) growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), doi:10.1073/pnas.0702737104.
Canadell, J. G., et al. (2010), Interactions of the carbon cycle, human activity, and the climate system: a research portfolio, *Current Opinion in Environmental Sustainability*, *2*(4), 301-311, doi:10.1016/j.cosust.2010.08.003.
Canadell J.G., P. Ciais, K. Gurney, C. Le Quéré, S. Piao, M. R. Raupach, C. L. Sabine (2011), An international effort to quantify regional carbon fluxes, *Eos Trans. AGU*, *92*(10), 81-82, doi:10.1029/2011EO100001

Cardinali, C. (2009), Monitoring the observation impact on the short-range forecast, *Quarterly Journal of the Royal Meteorological Society*, *135*(638), doi:10.1002/qj.366.
Cardinali, C., S. Pezzulli, and E. Andersson (2004), Influence-matrix diagnostic of a data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *130*(603), doi:10.1025/qj.03.205.
Carmichael, G. R., A. Sandu, T. Chai, D. N. Daescu, E. M. Constantinescu, and Y. Tang (2008), Predicting air quality: Improvements through advanced methods to integrate models and measurements, *Journal of Computational Physics*, *227*(7), 3540-3571, doi:10.1016/j.jcp.2007.02.024.
Carouge, C., P. Bousquet, P. Peylin, P. J. Rayner, and P. Ciais (2010a), What can we learn from European continuous atmospheric CO2 measurements to quantify regional fluxes - Part 1: Potential of the 2001 network, *Atmospheric Chemistry and Physics*, *10*(6), doi:10.5194/acp-10-3107-2010
Carouge, C., P. J. Rayner, P. Peylin, P. Bousquet, F. Chevallier, and P. Ciais (2010b), What can we learn from European continuous atmospheric CO2 measurements to quantify regional fluxes - Part 2: Sensitivity of flux accuracy to inverse setup, *Atmospheric Chemistry and Physics*, *10*(6), doi:10.5194/acp-10-3119-2010
Caya, A., J. Sun, and C. Snyder (2005), A comparison between the 4DVAR and the ensemble Kalman filter techniques for radar data assimilation, *Monthly Weather Review*, *133*(11), 3081-3094, doi:10.1175/mwr3021.1.
Chahine, M. T., et al. (2006), Improving weather forecasting and providing new data on greenhouse gases, *Bulletin of the American Meteorological Society*, *87*(7), 911-926, doi:10.1175/bams-87-7-911.
Chamberlin, T. C. (1897), A Group of Hypotheses Bearing on Climatic Changes, *J. Geology*, 5, 653-83.
Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand (2004), Properties and first application of an error-statistics tuning method in variational assimilation, *Quarterly Journal of the Royal Meteorological Society*, *130*(601), 2253-2275, doi:10.1256/qj.03.26.
Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand (2006), Diagnosis and tuning of observational error in a quasi-operational data assimilation setting, *Quarterly Journal of the Royal Meteorological Society*, *132*(615), doi:10.1256/qj.04.102.
Chatterjee, A., A. M. Michalak, J. L. Anderson, K. L. Mueller, V. Yadav (in press), Towards reliable ensemble Kalman filter estimates of CO2 fluxes, *Journal of Geophysical Research – Atmospheres*, doi:10.1029/2012JD018176
Chatterjee, A, R. J. Engelen, A . M. Michalak, S. R. Kawa, C. Sweeney (in prep.), Background error covariance for atmospheric CO2 data assimilation, in preparation for submission to *Journal of Geophysical Research-Atmospheres*
Chatterjee, A. and A. M. Michalak (in prep.), Inter-comparison of ensemble and variational data assimilation in the context of a CO2 source-sink estimation problem, in preparation for submission to *Geoscientific Model Development*.
Chedin, A., R. Saunders, A. Hollingsworth, N. Scott, M. Matricardi, J. Etcheto, C. Clerbaux, R. Armante, and C. Crevoisier (2003), The feasibility of monitoring CO2 from high-resolution infrared sounders, *Journal of Geophysical Research-Atmospheres*, *108*(D2), doi:10.1029/2001jd001443.
Cheng, H., M. Jardak, M. Alexe, and A. Sandu (2010), A hybrid approach to estimating error covariances in variational data assimilation, *Tellus Series a-Dynamic Meteorology and Oceanography*, *62*(3), doi:10.1111/j.1600-0870.2010.00442.x.
Chevallier, F., R. J. Engelen, and P. Peylin (2005a), The contribution of AIRS data to the estimation of CO(2) sources and sinks, *Geophysical Research Letters*, *32*(23), doi:10.1029/2005gl024229.
Chevallier, F., M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F. M. Breon, A. Chedin, and P. Ciais (2005b), Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, *Journal of Geophysical Research-Atmospheres*, *110*(D24), doi:10.1029/2005jd006390.
Chevallier, F., F.-M. Breon, and P. J. Rayner (2007), Contribution of the Orbiting Carbon Observatory to the estimation of CO2 sources and sinks: Theoretical study in a variational data assimilation framework, *Journal of Geophysical Research-Atmospheres*, *112*(D9), doi:10.1029/2006jd007375.
Chevallier, F., R. J. Engelen, C. Carouge, T. J. Conway, P. Peylin, C. Pickett-Heaps, M. Ramonet, P. J. Rayner, and I. Xueref-Remy (2009a), AIRS-based versus flask-based estimation of carbon surface

fluxes, *Journal of Geophysical Research-Atmospheres*, *114*, doi:10.1029/2009jd012311.
Chevallier, F., S. Maksyutov, P. Bousquet, F.-M. Breon, R. Saito, Y. Yoshida, and T. Yokota (2009b), On the accuracy of the CO2 surface fluxes to be estimated from the GOSAT observations, *Geophysical Research Letters*, *36*, doi:10.1029/2009gl040108.
Chevallier, F., et al. (2010a), CO2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *Journal of Geophysical Research-Atmospheres*, *115*, doi:10.1029/2010jd013887.
Chevallier, F., L. Feng, H. Boesch, P. I. Palmer, and P. J. Rayner (2010b), On the impact of transport model errors for the estimation of CO2 surface fluxes from GOSAT observations, *Geophysical Research Letters*, *37*, doi:10.1029/2010gl044652.
Chevallier, F., et al. (2011), Global CO2 fluxes inferred from surface air-sample measurements and from TCCON retrievals of the CO2 total column, *Geophysical Research Letters*, *38*, doi:10.1029/2011gl049899.
Chevallier, F., et al. (2012), What eddy-covariance measurements tell us about prior land flux errors in CO2-flux inversion schemes, *Global Biogeochemical Cycles*, *26*, doi:10.1029/2010gb003974.
Chiles, J.-P., and P. Delfiner (1999), Geostatistics: modeling spatial uncertainty, Wiley-Interscience, 695 pp.
Ciais, P., et al. (2010a), Can we reconcile atmospheric estimates of the Northern terrestrial carbon sink with land-based accounting?, *Current Opinion in Environmental Sustainability*, *2*(4), 225-230, doi:10.1016/j.cosust.2010.06.008.
Ciais, P., P. Rayner, F. Chevallier, P. Bousquet, M. Logan, P. Peylin, and M. Ramonet (2010b), Atmospheric inversions for estimating CO(2) fluxes: methods and perspectives, *Climatic Change*, *103*(1-2), 69-92, doi:10.1007/s10584-010-9909-3.
Cohn, S. E. (1997). An introduction to estimation theory: Data Assimilation in Meteorology and Oceanography: Theory and Practice, M. Ghil et al. (Eds.), *Universal Academic Publishers*, 211-218.
Cohn, S. E., A. da Silva, J. Guo, M. Sienkiewicz, and D. Lamich (1998), Assessing the effects of data selection with the DAO physical-space statistical analysis system, *Monthly Weather Review*, *126*(11), doi:10.1175/1520-0493(1998)126<2913:ateods>2.0.co;2.
Constantinescu, E. M., T. Chai, A. Sandu, and G. R. Carmichael (2007), Autoregressive models of background errors for chemical data assimilation, *Journal of Geophysical Research-Atmospheres*, *112*(D12), doi:10.1029/2006jd008103.
Conway, T.J., P.M. Lang, and K.A. Masarie (2011), Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2010, Version: 2011-10-14, ftp://ftp.cmdl.noaa.gov/ccg/co2/flask/event/.
Conway, T. J., P. P. Tans, L. S. Waterman, and K. W. Thoning (1994), Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration Climate Monitoring and Diagnostics Laboratory Global Air-Sampling Network, *Journal of Geophysical Research-Atmospheres*, *99*(D11), doi:10.1029/94jd01951.
Courtier, P. (1997), Dual formulation of four-dimensional variational assimilation, *Quarterly Journal of the Royal Meteorological Society*, *123*(544), 2449-2461, doi: 10.1002/qj.49712354414
Courtier, P., J. N. Thepaut, and A. Hollingsworth (1994), A Strategy for Operational Implementation of 4D-VAR using an Incremental Approach, *Quarterly Journal of the Royal Meteorological Society*, *120*(519), doi:10.1002/qj.49712051912.
Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell (2000), Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, *408*(6809), 184-187, doi:10.1038/35041539.
Cramer, W., et al. (2001), Global response of terrestrial ecosystem structure and function to CO2 and climate change: results from six dynamic global vegetation models, *Global Change Biology*, *7*(4), doi:10.1046/j.1365-2486.2001.00383.x.
Crevoisier, C., A. Chedin, H. Matsueda, T. Machida, R. Armante, and N. A. Scott (2009), First year of upper tropospheric integrated content of CO2 from IASI hyperspectral infrared observations,

Atmospheric Chemistry and Physics, *9*(14), 4797-4810, doi: 10.5194/acp-9-4797-2009
Crevoisier, C., C. Sweeney, M. Gloor, J. L. Sarmiento, and P. P. Tans (2010), Regional US carbon sinks from three-dimensional atmospheric CO2 sampling, *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18348-18353, doi:10.1073/pnas.0900062107.
Crisp, D., et al. (2004), The orbiting carbon observatory (OCO) mission, *Trace Constituents in the Troposphere and Lower Stratosphere*, *34*(4), doi:10.1016/j.asr.2003.08.062.
Crisp, D., et al. (2012), The ACOS CO2 retrieval algorithm - Part II: Global X-CO2 data characterization, *Atmospheric Measurement Techniques*, *5*(4), doi:10.5194/amt-5-687-2012.
Daley, R. (1991). Atmospheric Data Analysis, *Cambridge University Press*, 457 pp.
Daley, R., and E. Barker (2001), NAVDAS: Formulation and diagnostics, *Monthly Weather Review*, *129*(4), doi:10.1175/1520-0493(2001)129<0869:nfad>2.0.co;2.
da Silva, A., and J. Guo (1996) Documentation of the Physical-Space Statistical Analysis System (PSAS) Part I: The Conjugate Gradient Solver Version PSAS-1.00. DAO Office Note 96-02, 66 pp.
de Rosnay, P., J. Polcher, K. Laval, and M. Sabre (2003), Integrated parameterization of irrigation in the land surface model ORCHIDEE. Validation over Indian Peninsula, *Geophysical Research Letters*, *30*(19), doi:10.1029/2003gl018024.
Deckmyn, A., and L. Berre (2005), A wavelet approach to representing background error covariances in a limited-area model, *Monthly Weather Review*, *133*(5), 1279-1294, doi:10.1175/mwr2929.1.
Denman, K.L., G. Brasseur, A. Chidthaisong, P. Ciais, P.M. Cox, R.E. Dickinson, D. Hauglustaine, C. Heinze, E. Holland, D. Jacob, U. Lohmann, S Ramachandran, P.L. da Silva Dias, S.C. Wofsy and X. Zhang (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 (eds. Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Denning, A. S., G. J. Collatz, C. G. Zhang, D. A. Randall, J. A. Berry, P. J. Sellers, G. D. Colello, and D. A. Dazlich (1996a), Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model .1. Surface carbon fluxes, *Tellus Series B-Chemical and Physical Meteorology*, *48*(4), 521-542, doi:10.1034/j.1600-0889.1996.t01-2-00009.x.
Denning, A. S., D. A. Randall, G. J. Collatz, and P. J. Sellers (1996b), Simulations of terrestrial carbon metabolism and atmospheric CO2 in a general circulation model .2. Simulated CO2 concentrations, *Tellus Series B-Chemical and Physical Meteorology*, *48*(4), 543-567, doi:10.1034/j.1600-0889.1996.t01-1-00010.x.
Derber, J., and F. Bouttier (1999), A reformulation of the background error covariance in the ECMWF global data assimilation system, *Tellus Series a-Dynamic Meteorology and Oceanography*, *51*(2), 195-221, doi:10.1034/j.1600-0870.1999.t01-2-00003.x.
Desai, A. R. (2010), Climatic and phenological controls on coherent regional interannual variability of carbon dioxide flux in a heterogeneous landscape, *Journal of Geophysical Research-Biogeosciences*, *115*, 13, doi:10.1029/2010jg001423.
Desai, A. R., et al. (2011), Seasonal pattern of regional carbon balance in the central Rocky Mountains from surface and airborne measurements, *Journal of Geophysical Research-Biogeosciences*, *116*, 17, doi:10.1029/2011jg001655.
Desroziers, G., L. Berre, V. Chabot, and B. Chapnik (2009), A Posteriori Diagnostics in an Ensemble of Perturbed Analyses, *Monthly Weather Review*, *137*(10), doi:10.1175/2009mwr2778.1.
Di Giuseppe, F., C. Marsigli, and T. Paccagnella (2011), The relevance of background-error covariance matrix localization: an application to the variational retrieval of vertical profiles from SEVIRI observations, *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 29-42, doi:10.1002/qj.696.
Dietze, M. C., et al. (2011), Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American Carbon Program site-level synthesis, *Journal of Geophysical Research-Biogeosciences*, *116*, doi:10.1029/2011jg001661.

Dilling, L. (2007), Toward carbon governance: Challenges across scales in the United States, *Global Environmental Politics*, *7*(2), 28-44, doi:10.1162/glep.2007.7.2.28.

Dilling, L., S. C. Doney, J. Edmonds, K. R. Gurney, R. Harriss, D. Schimel, B. Stephens, and G. Stokes (2003), Therole of carbon cycle observations andknowledge in carbon management, *Annual Review of Environment and Resources*, *28*, doi:10.1146/annurev.energy.28.011503.163443.

Dirren, S., R. D. Torn, and G. J. Hakim (2007), A data assimilation case study using a limited-area ensemble Kalman filter, *Monthly Weather Review*, *135*(4), doi:10.1175/mwr3358.1.

Dolman, A. J., C. Gerbig, J. Noilhan, C. Sarrat, and F. Miglietta (2009), Detecting regional variability in sources and sinks of carbon dioxide: a synthesis, *Biogeosciences*, *6*(6), 1015-1026, doi:10.5194/bg-6-1015-2009

Ehrendorfer, M. (2007), A review of issues in ensemble-based Kalman filtering, *Meteorologische Zeitschrift*, *16*(6), 795-818, doi:10.1127/0941-2948/2007/0256.

El Akkraoui, A., P. Gauthier, S. Pellerin, and S. Buis (2008), Intercomparison of the primal and dual formulations of variational data assimilation, *Quarterly Journal of the Royal Meteorological Society*, *134*(633), doi:10.1002/qj.257.

Elbern, H., A. Strunk and L. Nieradzik (2010), Inverse Modeling and Combined State-Source Estimation for Chemical Weather, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 491-515, Springer-Verlag Berlin.

Eldering, A., S. Boland, B. Solish, D. Crisp, P. Kahn, and M. Gunson (2012), High precision atmospheric CO 2 measurements from space: the design and implementation of OCO-2, *2012 IEEE Aerospace Conference*, doi:10.1109/aero.2012.6187176.

Engelen, R. J., E. Andersson, F. Chevallier, A. Hollingsworth, M. Matricardi, A. P. McNally, J. N. Thepaut, and P. D. Watts (2004), Estimating atmospheric CO2 from advanced infrared satellite radiances within an operational 4D-Var data assimilation system: Methodology and first results, *Journal of Geophysical Research-Atmospheres*, *109*(D19), doi:10.1029/2004jd004777.

Engelen, R. J., A. S. Denning, K. R. Gurney, and M. TransCom (2002), On error estimation in atmospheric CO2 inversions, *Journal of Geophysical Research-Atmospheres*, *107*(D22), doi:10.1029/2002jd002195.

Engelen, R. J., and A. P. McNally (2005), Estimating atmospheric CO2 from advanced infrared satellite radiances within an operational four-dimensional variational (4D-Var) data assimilation system: Results and validation, *Journal of Geophysical Research-Atmospheres*, *110*(D18), doi:10.1029/2005jd005982.

Engelen, R. J., S. Serrar, and F. Chevallier (2009), Four-dimensional data assimilation of atmospheric CO2 using AIRS observations, *Journal of Geophysical Research-Atmospheres*, *114*, doi:10.1029/2008jd010739.

Engelen, R. J., and G. L. Stephens (2004), Information content of infrared satellite sounding measurements with respect to CO2, *Journal of Applied Meteorology*, *43*(2), 373-378, doi:10.1175/1520-0450(2004)043<0373:icoiss>2.0.co;2.

Enting, I. G., and J. V. Mansbridge (1989), Seasonal sources and sinks of atmospheric CO(2) Direct inversion of filtered data, *Tellus Series B-Chemical and Physical Meteorology*, *41*(2), doi:10.1111/j.1600-0889.1989.tb00129.x.

Enting, I.G. (2002). Inverse Problems in Atmospheric Constituent Transport. Atmospheric and Space Science Series, *Cambridge University Press*, Cambridge, 392 pp.

Errico, R. M. (1997), What is an adjoint model?, *Bulletin of the American Meteorological Society*, *78*(11), doi:10.1175/1520-0477(1997)078<2577:wiaam>2.0.co;2.

European Commission Report (2009) Integrated assessment of the European and North Atlantic Carbon Balance, Key results, policy implications for post 2012 and research needs, eds. Schulze, E.-D., C. Heinze, J. Gash, A. Volbers, A. Freibauer, A. Kentarchos, Luxembourg: Office for Official Publications of the European Communities, 137 pp.

Evensen, G. (1994), Sequential Data Assimilation with a Non-Linear Quasi-Geostrophic Model Using Monte-Carlo Methods to forecast Error Statistics, *Journal of Geophysical Research-Oceans*, *99*(C5),

doi:10.1029/94jc00572.
Evensen, G. (2003), The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics*, *53*(4), 343-367, doi: 10.1007/s10236-003-0036-9
Evensen, G. (2004), Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dynamics*, *54*(6), 539-560, doi:10.1007/s10236-004-0099-2.
Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson (1993), Assimilation of TOVS radiance information through one-dimensional variational analysis, *Quarterly Journal of the Royal Meteorological Society*, *119*(514), doi:10.1002/qj.49711951411.
Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans (1998), A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models, *Science*, *282*(5388), doi:10.1126/science.282.5388.442.
Feng, L., P. I. Palmer, H. Boesch, and S. Dance (2009), Estimating surface CO(2) fluxes from space-borne CO(2) dry air mole fraction observations using an ensemble Kalman Filter, *Atmospheric Chemistry and Physics*, *9*(8), 2619-2633, doi:10.5194/acp-9-2619-2009
Feng, L., P. I. Palmer, Y. Yang, R. M. Yantosca, S. R. Kawa, J. D. Paris, H. Matsueda, and T. Machida (2011), Evaluating a 3-D transport model of atmospheric CO2 using ground-based, aircraft, and space-borne data, *Atmospheric Chemistry and Physics*, *11*(6), doi:10.5194/acp-11-2789-2011.
Fertig, E. J., J. Harlim, and B. R. Hunt (2007), A comparative study of 4D-VAR and a 4D Ensemble Kalman Filter: perfect model simulations with Lorenz-96, *Tellus Series a-Dynamic Meteorology and Oceanography*, *59*(1), 96-100, doi:10.1111/j.1600-0870.2006.00205.x.
Finkelstein, P. L., and P. F. Sims (2001), Sampling error in eddy correlation flux measurements, *Journal of Geophysical Research-Atmospheres*, *106*(D4), 3503-3509, doi:10.1029/2000jd900731.
Fisher,M. (2004). Background error covariance modelling. Proc. of ECMWF Seminar on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, UK,45–64.
Fisher, M. (2006). "Wavelet"Jb –A new way to model the statistics of background errors. ECMWF Newsletter, 106, 23–28.Society, 130(601), 2253-2275, doi:10.1256/qj.03.26.
Fletcher, S. E. M., et al. (2007), Inverse estimates of the oceanic sources and sinks of natural CO2 and the implied oceanic carbon transport, *Global Biogeochemical Cycles*, *21*(1), doi:10.1029/2006gb002751.
Fourier, J (1824), Remarques Générales sur les Températures Du Globe Terrestre et des Espaces Planétaires, *Annales de Chemie et de Physique* 27, 136-67. Translation by Ebeneser Burgess, "General Remarks on the Temperature of the Earth and Outer Space," *American Journal of Science* 32, 1-20 (1837)
Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis: Results from the C(4)MIP model intercomparison, *Journal of Climate*, *19*(14), doi:10.1175/jcli3800.1.
Friedlingstein, P., R. A. Houghton, G. Marland, J. Hackler, T. A. Boden, T. J. Conway, J. G. Canadell, M. R. Raupach, P. Ciais, and C. Le Quere (2010), Update on CO2 emissions, *Nature Geoscience*, *3*(12), doi:10.1038/ngeo1022.
Friedlingstein, P., and I. C. Prentice (2010), Carbon-climate feedbacks: a review of model and observation based estimates, *Current Opinion in Environmental Sustainability*, *2*(4), doi:10.1016/j.cosust.2010.06.002.
Friend, A. D., et al. (2007), FLUXNET and modelling the global carbon cycle, *Global Change Biology*, *13*(3), 610-633, doi:10.1111/j.1365-2486.2006.01223.x.
Furrer, R., and T. Bengtsson (2007), Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants, *Journal of Multivariate Analysis*, *98*(2), 227-255, doi:10.1016/j.jmva.2006.08.003.
GCP (2009), Global Carbon Project - Carbon Budget 2008, Released on 17 November, 2009, Available at: http://www.globalcarbonproject.org/carbonbudget/index.htm
GLOBALVIEW-CO2 (2008), Cooperative Atmospheric Data Integration Project—Carbon dioxide [CD-ROM], Clim. Monit. and Diag. Lab., NOAA, Boulder, Colo. (also available via anonymous ftp to ftp.cmdl.noaa.gov, Path: ccg/CO2/GLOBALVIEW).
Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions,

Quarterly Journal of the Royal Meteorological Society, *125*(554), 723-757, doi:10.1256/smsqj.55416.
Geels, C., S. C. Doney, R. Dargaville, J. Brandt, and J. H. Christensen (2004), Investigating the sources of synoptic variability in atmospheric CO2 measurements over the Northern Hemisphere continents: a regional model study, *Tellus Series B-Chemical and Physical Meteorology*, *56*(1), doi:10.1111/j.1600-0889.2004.00084.x.
Geels, C., et al. (2007), Comparing atmospheric transport models for future regional inversions over Europe - Part 1: mapping the atmospheric CO2 signals, *Atmospheric Chemistry and Physics*, *7*(13), 3461-3479, doi:10.5194/acp-7-3461-2007
Gerber, M., and F. Joos (2010), Carbon sources and sinks from an Ensemble Kalman Filter ocean data assimilation, *Global Biogeochemical Cycles*, *24*, doi:10.1029/2009gb003531.
Gerbig, C., A. J. Dolman, and M. Heimann (2009), On observational and modelling strategies targeted at regional carbon exchange over continents, *Biogeosciences*, *6*(10), doi:10.5194/bg-6-1949-2009
Gerbig, C., S. Korner, and J. C. Lin (2008), Vertical mixing in atmospheric tracer transport models: error characterization and propagation, *Atmospheric Chemistry and Physics*, *8*(3), 591-602, doi:10.5194/acp-8-591-2008
Gerbig, C., J. C. Lin, J. W. Munger, and S. C. Wofsy (2006), What can tracer observations in the continental boundary layer tell us about surface-atmosphere fluxes?, *Atmospheric Chemistry and Physics*, *6*, 539-554, doi:10.5194/acp-6-539-2006
Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger (2003a), Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, *Journal of Geophysical Research-Atmospheres*, *108*(D24), doi:10.1029/2002jd003018.
Gerbig, C., J. C. Lin, S. C. Wofsy, B. C. Daube, A. E. Andrews, B. B. Stephens, P. S. Bakwin, and C. A. Grainger (2003b), Toward constraining regional-scale fluxes of CO2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *Journal of Geophysical Research-Atmospheres*, *108*(D24), 27, doi: 10.1029/2003jd003770.
Giering, R. (2000) Tangent Linear and Adjoint Biogeochemical Models, in *Inverse methods in global biogeochemical models*, eds. P. Kasibhatla et al., Geophysical monograph 114, American Geophysical Union, Washington DC.
Ghil, M., and P. Malanotterizzoli (1991), Data Assimilation in Meteorology and Oceanography, *Advances in Geophysics*, *33*, doi:10.1016/s0065-2687(08)60442-2.
Gloor, M., P. Bakwin, D. Hurst, L. Lock, R. Draxler, and P. Tans (2001), What is the concentration footprint of a tall tower?, *Journal of Geophysical Research-Atmospheres*, *106*(D16), 17831-17840, doi:10.1029/2001jd900021.
Gloor, M., S. M. Fan, S. Pacala, J. Sarmiento, and M. Ramonet (1999), A model-based evaluation of inversions of atmospheric transport, using annual mean mixing ratios, as a tool to monitor fluxes of nonreactive trace substances like CO2 on a continental scale, *Journal of Geophysical Research-Atmospheres*, *104*(D12), doi:10.1029/1999jd900132.
Gloor, M., J. L. Sarmiento, and N. Gruber (2010), What can be learned about carbon cycle climate feedbacks from the CO2 airborne fraction?, *Atmospheric Chemistry and Physics*, *10*(16), 7739-7751, doi:10.5194/acp-10-7739-2010.
Gneiting, T. (2002), Compactly supported correlation functions, *Journal of Multivariate Analysis*, *83*(2), 493-508, doi:10.1006/jmva.2001.2056.
Gockede, M., A. M. Michalak, D. Vickers, D. P. Turner, and B. E. Law (2010), Atmospheric inverse modeling to constrain regional-scale CO2 budgets at high spatial and temporal resolution, *Journal of Geophysical Research-Atmospheres*, *115*, doi:10.1029/2009jd012257.
Goetz, S. J., A. Baccini, N. T. Laporte, T. Johns, W. Walker, J. Kellndorfer, R. A. Houghton, and M. Sun (2009), Mapping and monitoring carbon stocks with satellite observations: a comparison of methods, *Carbon Balance and Management*, *4*(2), doi:10.1186/1750-0680-4-2
Gourdji, S. M., A. I. Hirsch, K. L. Mueller, V. Yadav, A. E. Andrews, and A. M. Michalak (2010), Regional-scale geostatistical inverse modeling of North American CO(2) fluxes: a synthetic data

study, *Atmospheric Chemistry and Physics*, *10*(13), 6151-6167, doi:10.5194/acp-10-6151-2010.
Gourdji, S. M., K. L. Mueller, K. Schaefer, and A. M. Michalak (2008), Global monthly averaged CO(2) fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *Journal of Geophysical Research-Atmospheres*, *113*(D21), doi:10.1029/2007jd009733.
Gourdji, S. M., et al. (2012), North American CO2 exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion, *Biogeosciences*, *9*(1), doi:10.5194/bg-9-457-2012.
Gratton, S., and J. Tshimanga (2009), An observation-space formulation of variational assimilation using a restricted preconditioned conjugate gradient algorithm, *Quarterly Journal of the Royal Meteorological Society*, *135*(643), doi:10.1002/qj.477.
Gruber, N., et al. (2009), Oceanic sources, sinks, and transport of atmospheric CO2, *Global Biogeochemical Cycles*, *23*, 21, doi:10.1029/2008gb003349.
Gurney, K. R., and W. J. Eckels (2011), Regional trends in terrestrial carbon exchange and their seasonal signatures, *Tellus Series B-Chemical and Physical Meteorology*, *63*(3), 328-339, doi:10.1111/j.1600-0889.2011.00534.x.
Gurney, K. R., et al. (2003), TransCom 3 CO2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus Series B-Chemical and Physical Meteorology*, *55*(2), 555-579 doi: 10.1034/j.1600-0889.2003.00049.x
Gurney, K. R., et al. (2004), Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochemical Cycles*, *18*(1), doi:10.1029/2003gb002111.
Gurney, K. R., D. L. Mendoza, Y. Zhou, M. L. Fischer, C. C. Miller, S. Geethakumar, and S. D. L. R. Du Can (2009), High Resolution Fossil Fuel Combustion CO2 Emission Fluxes for the United States, *Environmental Science & Technology*, *43*(14), doi:10.1021/es900806c.
Haines, K. (2010), Ocean Data Assimilation, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 517-547, Springer-Verlag Berlin.
Hamazaki, T., Y. Kaneko, A. Kuze, and K. Kondo (2005), Fourier transform spectrometer for Greenhouse gases Observing Satellite (GOSAT), paper presented at Conference on Enabling Sensor and Platform Technologies for Spaceborne Remote Sensing, Honolulu, HI, Nov 09-10.
Hamazaki, T., A. Kuze, and K. Kondo (2004), Sensor system for greenhouse gas observing satellite (GOSAT), paper presented at 12th Conference on Infrared Spaceborne Remote Sensing, Spie-Int Soc Optical Engineering, Denver, CO, Aug 02-03.
Hamill, T. M., and J. S. Whitaker (2005), Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches, *Monthly Weather Review*, *133*(11), 3132-3147, doi:10.1175/mwr3020.1.
Hamill, T. M., and J. S. Whitaker (2011), What Constrains Spread Growth in Forecasts Initialized from Ensemble Kalman Filters?, *Monthly Weather Review*, *139*(1), 117-131, doi:10.1175/2010mwr3246.1.
Hamill, T. M., J. S. Whitaker, and C. Snyder (2001), Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Monthly Weather Review*, *129*(11), 2776-2790, doi:10.1175/1520-0493(2001)129<2776:ddfobe>2.0.co;2.
Hammerling, D. M., A. M. Michalak, and S. R. Kawa (2012a), Mapping of CO2 at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, *Journal of Geophysical Research-Atmospheres*, *117*, doi:10.1029/2011jd017015.
Hammerling, D. M., A. M. Michalak, C. O'Dell, and S. R. Kawa (2012b), Global CO2 distributions over land from the Greenhouse Gases Observing Satellite (GOSAT), *Geophysical Research Letters*, *39*, 6, doi:10.1029/2012gl051203.
Hayes, D. J., et al. (2012), Reconciling estimates of the contemporary North American carbon balance among terrestrial biosphere models, atmospheric inversions, and a new approach for estimating net ecosystem exchange from inventory-based data, *Global Change Biology*, *18*(4), doi:10.1111/j.1365-2486.2011.02627.x.
Heimann, M., et al. (1998), Evaluation of terrestrial Carbon Cycle models through simulations of the

seasonal cycle of atmospheric CO(2): First results of a model intercomparison study, *Global Biogeochemical Cycles*, *12*(1), 1-24, doi:10.1029/97gb01936.
Heimann, M., and M. Reichstein (2008), Terrestrial ecosystem carbon dynamics and climate feedbacks, *Nature*, *451*(7176), doi:10.1038/nature06591.
He, H. L., et al. (2010), Uncertainty analysis of eddy flux measurements in typical ecosystems of ChinaFLUX, *Ecological Informatics*, *5*(6), 492-502, doi:10.1016/j.ecoinf.2010.07.004.
Henze, D. K., A. Hakami, and J. H. Seinfeld (2007), Development of the adjoint of GEOS-Chem, *Atmospheric Chemistry and Physics*, *7*(9), 2413-2433, doi:10.5194/acp-7-2413-2007
Hess, R. (2010), Flow dependence of background errors and their vertical correlations for radiance-data assimilation, *Quarterly Journal of the Royal Meteorological Society*, *136*(647), 475-486, doi:10.1002/qj.570.
Hofmann, D. J., J. H. Butler, E. J. Dlugokencky, J. W. Elkins, K. Masarie, S. A. Montzka, and P. Tans (2006), The role of carbon dioxide in climate forcing from 1979 to 2004: introduction of the Annual Greenhouse Gas Index, *Tellus Series B-Chemical and Physical Meteorology*, *58*(5), doi:10.1111/j.1600-0889.2006.00201.x.
Hollinger, D. Y., and A. D. Richardson (2005), Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiology*, *25*(7), 873-885, doi: 10.1093/treephys/25.7.873
Hollingsworth, A., et al. (2008), Toward a monitoring and forecasting system for atmospheric composition: The GEMS project, *Bulletin of the American Meteorological Society*, *89*(8), 1147-1164, doi:10.1175/2008bams2355.1.
Horn, R. A., and R. Mathias (1990), An analog of the Cauchy-Schwarz Inequality for Hadamard products and Unitarily Invariant Norms, *Siam Journal on Matrix Analysis and Applications*, *11*(4), 481-498, doi:10.1137/0611034.
Houghton, R.A., and Hackler J. L. (2001) *Carbon Flux to the Atmosphere From Land-use Changes: 1850 to 1990*. NDP-050/R1, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee.
Houghton, R. A. (2001), Counting terrestrial sources and sinks of carbon, *Climatic Change*, *48*(4), doi:10.1023/a:1005658316062.
Houghton, R. A. (2003), Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850-2000, *Tellus Series B-Chemical and Physical Meteorology*, *55*(2), doi:10.1034/j.1600-0889.2003.01450.x.
Houghton, R. A. (2005), Aboveground forest biomass and the global carbon balance, *Global Change Biology*, *11*(6), 945-958, doi:10.1111/j.1365-2486.2005.00955.x.
Houghton, R. A. (2007), Balancing the global carbon budget, in *Annual Review of Earth and Planetary Sciences*, edited, pp. 313-347, Annual Reviews, Palo Alto, doi:10.1146/annurev.earth.35.031306.140057.
Houghton, R. A. (2010), How well do we know the flux of CO2 from land-use change?, *Tellus Series B-Chemical and Physical Meteorology*, *62*(5), doi:10.1111/j.1600-0889.2010.00473.x.
Houghton, R. A., M. Gloor, J. Lloyd, and C. Potter (2009), The Regional Carbon Budget, *Amazonia and Global Change*, *186*, doi:10.1029/2008gm000718.
Houser, P. R, G. J. M De Lannoy and J. P. Walker. (2010), Land Surface Data Assimilation, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 549-597, Springer-Verlag Berlin
Houtekamer, P. L., and H. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Monthly Weather Review*, *126*(3), 796-811, doi:10.1175/1520-0493(1998)126<0796:dauaek>2.0.co;2.
Houtekamer, P. L., and H. L. Mitchell (2001), A sequential ensemble Kalman filter for atmospheric data assimilation, *Monthly Weather Review*, *129*(1), 123-137, doi:10.1175/1520-0493(2001)129<0123:asekff>2.0.co;2.
Houtekamer, P. L., and H. L. Mitchell (2005), Ensemble Kalman filtering, *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3269-3289, doi:10.1256/qj.05.135.

Houtekamer, P. L., H. L. Mitchell, and X. Deng (2009), Model Error Representation in an Operational Ensemble Kalman Filter, *Monthly Weather Review*, *137*(7), 2126-2143, doi:10.1175/2008mwr2737.1.
Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and M. Hansen (2005), Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations, *Monthly Weather Review*, *133*(3), 604-620, doi:10.1175/mwr-2864.1.
Houweling, S., et al. (2010), The importance of transport model uncertainties for the estimation of CO2 sources and sinks using satellite measurements, *Atmospheric Chemistry and Physics*, *10*(20), doi:10.5194/acp-10-9981-2010.
Houweling, S., F. M. Breon, I. Aben, C. Rodenbeck, M. Gloor, M. Heimann, and P. Ciais (2004), Inverse modeling of CO2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time, *Atmospheric Chemistry and Physics*, *4*, 523-538, doi:10.5194/acp-4-523-2004
Hungershoefer, K., F. M. Breon, P. Peylin, F. Chevallier, P. Rayner, A. Klonecki, S. Houweling, and J. Marshall (2010), Evaluation of various observing systems for the global monitoring of CO2 surface fluxes, *Atmospheric Chemistry and Physics*, *10*(21), doi:10.5194/acp-10-10503-2010.
Hunt, B. R., E. J. Kostelich, and I. Szunyogh (2007), Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D-Nonlinear Phenomena*, *230*(1-2), doi:10.1016/j.physd.2006.11.008.
Huntzinger, D. N., S. M. Gourdji, K. L. Mueller, and A. M. Michalak (2011), The utility of continuous atmospheric measurements for identifying biospheric CO2 flux variability, *Journal of Geophysical Research-Atmospheres*, *116*, doi:10.1029/2010jd015048.
Huntzinger, D. N., et al. (2012), North American Carbon Program (NACP) regional interim synthesis: Terrestrial biospheric model intercomparison, *Ecological Modelling*, *232*, doi:10.1016/j.ecolmodel.2012.02.004.
IPCC (2007), Climate Change 2007:Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, Pachauri, R. K. and Reisinger, A. (eds.)], IPCC, Geneva, Switzerland, 104 pp.
Janiskova, M., P. Lopez, and P. Bauer (2012), Experimental 1D+4D-Var assimilation of CloudSat observations, *Quarterly Journal of the Royal Meteorological Society*, *138*(666), doi:10.1002/qj.988.
Jardak, M., I. M. Navon, and M. Zupanski (2010), Comparison of sequential data assimilation methods for the Kuramoto-Sivashinsky equation, *International Journal for Numerical Methods in Fluids*, *62*(4), 374-402, doi:10.1002/fld.2020.
Jazwinski, Andrew H. (1970). Stochastic Processes and Filtering. Mathematics in Science and Engineering. New York: Academic Press. pp. 376
Jun, M., I. Szunyogh, M. G. Genton, F. Zhang, and C. H. Bishop (2011), A Statistical Investigation of the Sensitivity of Ensemble-Based Kalman Filters to Covariance Filtering, *Monthly Weather Review*, *139*(9), doi:10.1175/2011mwr3577.1.
Jung, M., et al. (2011), Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *Journal of Geophysical Research - Biogeosciences*, 116, G00J07, doi:10.1029/2010JG001566
Kadygrov, N., S. Maksyutov, N. Eguchi, T. Aoki, T. Nakazawa, T. Yokota, and G. Inoue (2009), Role of simulated GOSAT total column CO2 observations in surface CO2 flux uncertainty reduction, *Journal of Geophysical Research-Atmospheres*, *114*, doi:10.1029/2008jd011597.
Kahnert, M. (2008), Variational data analysis of aerosol species in a regional CTM: background error covariance constraint and aerosol optical observation operators, *Tellus Series B-Chemical and Physical Meteorology*, *60*(5), 753-770, doi:10.1111/j.1600-0889.2008.00377.x.
Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Transaction of the ASME—Journal of Basic Engineering*, *82*, Series D, 35-45.
Kalnay, E. (2003) Atmospheric Modeling, Data Assimilation and Predictability. *Cambridge University Press*, Cambridge, 341 pp.
Kalnay, E., H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy (2007), 4-D-Var or ensemble Kalman

filter?, *Tellus Series a-Dynamic Meteorology and Oceanography*, *59*(5), 758-773, doi:10.1111/j.1600-0870.2007.00261.x.
Kaminski, T., M. Heimann, and R. Giering (1999), A coarse grid three-dimensional global inverse model of the atmospheric transport - 2. Inversion of the transport of CO(2) in the 1980s, *Journal of Geophysical Research-Atmospheres*, *104*(D15), 18555-18581, doi:10.1029/1999jd900146.
Kaminski, T., W. Knorr, P. J. Rayner, and M. Heimann (2002), Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, *Global Biogeochemical Cycles*, *16*(4), doi:10.1029/2001gb001463.
Kaminski, T., P. J. Rayner, M. Heimann, and I. G. Enting (2001), On aggregation errors in atmospheric transport inversions, *Journal of Geophysical Research-Atmospheres*, *106*(D5), 4703-4715, doi:10.1029/2000jd900581.
Kaminski, T.,P. J. Rayner, M. Voßbeck, M. Scholze, and E. Koffi (2012) Observing the continental-scale carbon balance: assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design. *Atmospheric Chemistry and Physics*, 12(16):7867-7879, doi:10.5194/acp-12-7867-2012
Kang, J.-S., E. Kalnay, J. Liu, I. Fung, T. Miyoshi, and K. Ide (2011), "Variable localization" in an ensemble Kalman filter: Application to the carbon cycle data assimilation, *Journal of Geophysical Research-Atmospheres*, *116*, doi:10.1029/2010jd014673.
Karspeck, A. R., and J. L. Anderson (2007), Experimental implementation of an ensemble adjustment filter for an intermediate ENSO model, *Journal of Climate*, *20*(18), 4638-4658, doi:10.1175/jcli4245.1.
Kawa, S. R., D. J. Erickson, S. Pawson, and Z. Zhu (2004), Global CO2 transport simulations using meteorological data from the NASA data assimilation system, *Journal of Geophysical Research-Atmospheres*, *109*(D18), doi:10.1029/2004jd004554.
Kawa, S. R., J. Mao, J. B. Abshire, G. J. Collatz, X. Sun, and C. J. Weaver (2010), Simulation studies for a space-based CO2 lidar mission, *Tellus Series B-Chemical and Physical Meteorology*, *62*(5), 759-769, doi:10.1111/j.1600-0889.2010.00486.x.
Keeling, R. F., S. C. Piper, and M. Heimann (1996), Global and hemispheric CO2 sinks deduced from changes in atmospheric O-2 concentration, *Nature*, *381*(6579), 218-221, doi:10.1038/381218a0.
Keenan, T. F., et al. (2012), Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO2 exchange, *Global Change Biology*, *18*(6), 1971-1987, doi:10.1111/j.1365-2486.2012.02678.x.
Keenan, T. F., M. S. Carbone, M. Reichstein, and A. D. Richardson (2011), The model-data fusion pitfall: assuming certainty in an uncertain world, *Oecologia*, *167*(3), doi:10.1007/s00442-011-2106-x.
Kepert, J. D. (2009), Covariance localisation and balance in an Ensemble Kalman Filter, *Quarterly Journal of the Royal Meteorological Society*, *135*(642), 1157-1176, doi:10.1002/qj.443.
Kepert, J. D. (2011), Balance-aware covariance localisation for atmospheric and oceanic ensemble Kalman filters, *Computational Geosciences*, *15*(2), 239-250, doi:10.1007/s10596-010-9188-0.
King, D. A., D. P. Turner, and W. D. Ritts (2011), Parameterization of a diagnostic carbon cycle model for continental scale application, *Remote Sensing of Environment*, *115*(7), 1653-1664, doi:10.1016/j.rse.2011.02.024.
Kitanidis, P. K. (1995), Quasi-Linear Geostatistical Theory for Inversing, W*ater Resources Research*, *31*(10), 2411-2419, doi:10.1029/95wr01945.
Kitanidis, P. K. (1997), A variance-ratio test for supporting a variable mean in kriging, *Mathematical Geology*, *29*(3), doi:10.1007/bf02769639.
Kitanidis, P. K., and E. G. Vomvoris (1983), A geostatistical Approach to the Inverse problem in Groundwater Modeling (Steady-State) and one-dimensional simulations, *Water Resources Research*, *19*(3), doi:10.1029/WR019i003p00677.
Kleiman, G., and R. G. Prinn (2000), Measurement and deduction of emissions of trichloroethene, tetrachloroethene, and trichloromethane (chloroform) in the northeastern United States and southeastern Canada, *Journal of Geophysical Research-Atmospheres*, *105*(D23), 28875-28893,

doi:10.1029/2000jd900513.
Knorr, W., T. Kaminski, M. Scholze, N. Gobron, B. Pinty, R. Giering, and P. P. Mathieu (2010), Carbon cycle data assimilation with a generic phenology model, *Journal of Geophysical Research-Biogeosciences*, *115*, doi:10.1029/2009jg001119.
Koffi, E. N., P. J. Rayner, M. Scholze, and C. Beer (2012), Atmospheric constraints on gross primary productivity and net ecosystem productivity: Results from a carbon-cycle data assimilation system, *Global Biogeochemical Cycles*, *26*, doi:10.1029/2010gb003900.
Kretschmer, R., C. Gerbig, U. Karstens, and F. T. Koch (2012), Error characterization of CO2 vertical mixing in the atmospheric transport model WRF-VPRM, *Atmospheric Chemistry and Physics*, *12*(5), doi:10.5194/acp-12-2441-2012.
Krinner, G., N. Viovy, N. de Noblet-Ducoudre, J. Ogee, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochemical Cycles*, *19*(1), doi:10.1029/2003gb002199.
Krol, M., S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener, and P. Bergamaschi (2005), The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, *Atmospheric Chemistry and Physics*, *5*, 417-432, doi:10.5194/acp-5-417-2005
Kulawik, S. S., et al. (2010), Characterization of Tropospheric Emission Spectrometer (TES) CO2 for carbon cycle science, *Atmospheric Chemistry and Physics*, *10*(12), 5601-5623, doi:10.5194/acp-10-5601-2010.
Kurapov, A. L., G. D. Egbert, R. N. Miller, and J. S. Allen (2002), Data assimilation in a baroclinic coastal ocean model: Ensemble statistics and comparison of methods, *Monthly Weather Review*, *130*(4), 1009-1025, doi:10.1175/1520-0493(2002)130<1009:daiabc>2.0.co;2.
Kuze, A., H. Suto, M. Nakajima, and T. Hamazaki (2009), Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Applied Optics*, *48*(35), 6716-6733, doi: 10.1364/AO.48.006716
Lahoz, W. A., Q. Errera, R. Swinbank, and D. Fonteyn (2007), Data assimilation of stratospheric constituents: a review, *Atmospheric Chemistry and Physics*, *7*(22), 5745-5773, doi:10.5194/acp-7-5745-2007
Lahoz, W and Q. Errera. (2010), Constituent Assimilation, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 449-489, Springer-Verlag Berlin.
Lahoz, W., B. Khattativ, R. Menard (eds.) (2010) Data Assimilation: Making Sense of Observations, Springer-Verlag Berlin Heidelberg 2010, 713 pp, doi 10.1007/978-3-540-74703-1
Lasslop, G., M. Reichstein, J. Kattge, and D. Papale (2008), Influences of observation errors in eddy flux data on inverse model parameter estimation, *Biogeosciences*, *5*(5), 1311-1324, doi:10.5194/bg-5-1311-2008
Lauvaux, T., et al. (2009), Bridging the gap between atmospheric concentrations and local ecosystem measurements, *Geophysical Research Letters*, *36*, doi:10.1029/2009gl039574.
Lauvaux, T., et al. (2012), Constraining the CO2 budget of the corn belt: exploring uncertainties from the assumptions in a mesoscale inverse system, *Atmospheric Chemistry and Physics*, *12*(1), doi:10.5194/acp-12-337-2012.
Law, R. M., Y. H. Chen, K. R. Gurney, and M. TransCom (2003), TransCom 3 CO2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, *Tellus Series B-Chemical and Physical Meteorology*, *55*(2), doi:10.1034/j.1600-0889.2003.00053.x.
Law, R. M., L. P. Steele, P. B. Krummel, and W. Zahorowski (2010), Synoptic variations in atmospheric CO2 at Cape Grim: a model intercomparison, *Tellus Series B-Chemical and Physical Meteorology*, *62*(5), 810-820, doi:10.1111/j.1600-0889.2010.00470.x.
Lawson, W. G., and J. A. Hansen (2004), Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth, *Monthly Weather Review*, *132*(8), 1966-1981, doi:10.1175/1520-0493(2004)132<1966:iosadf>2.0.co;2.
Le Dimet, F. X., I. M. Navon, and D. N. Daescu (2002), Second-order information in data assimilation,

Monthly Weather Review, *130*(3), doi:10.1175/1520-0493(2002)130<0629:soiida>2.0.co;2.
Le Quere, C., et al. (2009), Trends in the sources and sinks of carbon dioxide, *Nature Geoscience*, *2*(12), doi:10.1038/ngeo689.
Ledimet, F. X., and O. Talagrand (1986), Variational Algorithms for Analysis and Assimilation of Meteorological observations – theoretical aspects, *Tellus Series a-Dynamic Meteorology and Oceanography*, *38*(2), 97-110, doi:10.1111/j.1600-0870.1986.tb00459.x
Leeuwenburgh, O., G. Evensen, and L. Bertino (2005), The impact of ensemble filter definition on the assimilation of temperature profiles in the tropical Pacific, *Quarterly Journal of the Royal Meteorological Society*, *131*(613), doi:10.1256/qj.05.90.
Lenton, T. M., H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber (2008), Tipping elements in the Earth's climate system, *Proceedings of the National Academy of Sciences of the United States of America*, *105*(6), doi:10.1073/pnas.0705414105.
Levin, I., and U. Karstens (2007), Inferring high-resolution fossil fuel CO2 records at continental sites from combined (CO2)-C-14 and CO observations, *Tellus Series B-Chemical and Physical Meteorology*, *59*(2), doi:10.1111/j.1600-0889.2006.00244.x.
Lewis, J.M., S. Lakshmivarahan, and S. Dhall, (2006). Dynamic Data Assimilation: A Least Squares Approach, *Cambridge University Press*, Cambridge, 2006.
Li, H., E. Kalnay, and T. Miyoshi (2009), Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, *Quarterly Journal of the Royal Meteorological Society*, *135*(639), 523-533, doi:10.1002/qj.371.
Lin, J. C., C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger (2003), A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *Journal of Geophysical Research-Atmospheres*, *108*(D16), 18, doi: 10.1029/2002jd003161.
Lin, S. J. (2004), A "vertically Lagrangian" finite-volume dynamical core for global models, *Monthly Weather Review*, *132*(10), doi:10.1175/1520-0493(2004)132<2293:avlfdc>2.0.co;2.
Liu, J., E. J. Fertig, H. Li, E. Kalnay, B. R. Hunt, E. J. Kostelich, I. Szunyogh, and R. Todling (2008), Comparison between Local Ensemble Transform Kalman Filter and PSAS in the NASA finite volume GCM - perfect model experiments, *Nonlinear Processes in Geophysics*, *15*(4), 645-659, doi:10.5194/npg-15-645-2008
Liu, J., I. Fung, E. Kalnay, and J.-S. Kang (2011), CO(2) transport uncertainties from the uncertainties in meteorological fields, *Geophysical Research Letters*, *38*, doi: 10.1029/2011gl047213.
Liu, J., I. Fung, E. Kalnay, J.-S. Kang, E. T. Olsen, and L. Chen (2012), Simultaneous assimilation of AIRS Xco(2) and meteorological observations in a carbon climate model with an ensemble Kalman filter, *Journal of Geophysical Research-Atmospheres*, *117*, doi:10.1029/2011jd016642.
Liu, J., E. Kalnay, T. Miyoshi, and C. Cardinali (2009), Analysis sensitivity calculation in an ensemble Kalman filter, *Quarterly Journal of the Royal Meteorological Society*, *135*(644), doi:10.1002/qj.511.
Loescher, H. W., B. E. Law, L. Mahrt, D. Y. Hollinger, J. Campbell, and S. C. Wofsy (2006), Uncertainties in, and interpretation of, carbon flux estimates using the eddy covariance technique, *Journal of Geophysical Research-Atmospheres*, *111*(D21), 19, doi:10.1029/2005jd006932.
Lokupitiya, R. S., D. Zupanski, A. S. Denning, S. R. Kawa, K. R. Gurney, and M. Zupanski (2008), Estimation of global CO(2) fluxes at regional scale using the maximum likelihood ensemble filter, *Journal of Geophysical Research-Atmospheres*, *113*(D20), doi: 10.1029/2007jd009679.
Lorenc, A. C. (1986), Analysis Methods for Numerical Weather Prediction, *Quarterly Journal of the Royal Meteorological Society*, *112*(474), doi:10.1002/qj.49711247414.
Lorenc, A. C. (2003), The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var, *Quarterly Journal of the Royal Meteorological Society*, *129*(595), 3183-3203, doi:10.1256/qj.02.132.
Lupu, C., P. Gauthier, and S. Laroche (2012), Assessment of the Impact of Observations on Analyses Derived from Observing System Experiments, *Monthly Weather Review*, *140*(1), 245-257, doi:10.1175/mwr-d-10-05010.1.
Luyssaert, S., et al. (2009), Toward a consistency cross-check of eddy covariance flux-based and

biometric estimates of ecosystem carbon balance, *Global Biogeochemical Cycles*, *23*, 13, doi:10.1029/2008gb003377.
Machida, T., H. Matsueda, Y. Sawa, Y. Nakagawa, K. Hirotani, N. Kondo, K. Goto, T. Nakazawa, K. Ishikawa, and T. Ogawa (2008), Worldwide Measurements of Atmospheric CO2 and Other Trace Gas Species Using Commercial Airlines, *Journal of Atmospheric and Oceanic Technology*, *25*(10), doi:10.1175/2008jtecha1082.1.
Maksyutov, S., N. Kadygrov, Y. Nakatsuka, P. K. Patra, T. Nakazawa, T. Yokota, and G. Inoue (2008), Projected impact of the GOSAT observations on regional CO2 flux estimations as a function of total retrieval error, J. Remote Sens. Soc. Jpn., 28, 190– 197.
Manning, A. C., and R. F. Keeling (2006), Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network, *Tellus Series B-Chemical and Physical Meteorology*, *58*(2), doi:10.1111/j.1600-0889.2006.00175.x.
Manning, A. C., E. G. Nisbet, R. F. Keeling, and P. S. Liss (2011), Greenhouse gases in the Earth system: setting the agenda to 2030, *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, *369*(1943), 1885-1890, doi:10.1098/rsta.2011.0076.
Manning, A. J. (2011), The challenge of estimating regional trace gas emissions from atmospheric observations, *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, *369*(1943), doi:10.1098/rsta.2010.0321.
Marecal, V., and J. F. Mahfouf (2002), Four-dimensional variational assimilation of total column water vapor in rainy areas, *Monthly Weather Review*, *130*(1), 43-58, doi:10.1175/1520-0493(2002)130<0043:fdvaot>2.0.co;2.
Masarie, K. A., and P. P. Tans (1995), Extension and Integration of Atmospheric Carbon Dioxide Data into a Globally Consistent Measurement Record, *Journal of Geophysical Research-Atmospheres*, *100*(D6), 11593-11610, doi:10.1029/95jd00859.
Massart, S., B. Pajot, A. Piacentini, and O. Pannekoucke (2010), On the Merits of Using a 3D-FGAT Assimilation Scheme with an Outer Loop for Atmospheric Situations Governed by Transport, *Monthly Weather Review*, *138*(12), doi:10.1175/2010mwr3237.1.
Massart, S., A. Piacentini, and O. Pannekoucke (2012), Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses, *Quarterly Journal of the Royal Meteorological Society*, *138*(665), 889-905, doi:10.1002/qj.971.
Matross, D. M., et al. (2006), Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, *Tellus Series B-Chemical and Physical Meteorology*, *58*(5), 344-358, doi:10.1111/j.1600-0889.2006.00206.x.
Matsumoto, K., et al. (2004), Evaluation of ocean carbon cycle models with data-based metrics, *Geophysical Research Letters*, *31*(7), 4, doi:10.1029/2003gl018970.
Matthews, H. D., M. Eby, T. Ewen, P. Friedlingstein, and B. J. Hawkins (2007), What determines the magnitude of carbon cycle-climate feedbacks?, *Global Biogeochemical Cycles*, *21*(2), 12, doi:10.1029/2006gb002733.
McKinley, D. C., et al. (2011), A synthesis of current knowledge on forests and carbon storage in the United States, *Ecological Applications*, *21*(6), 1902-1924, doi:10.1890/10-0697.1.
Medvigy, D., and P. R. Moorcroft (2012), Predicting ecosystem dynamics at regional scales: an evaluation of a terrestrial biosphere model for the forests of northeastern North America, *Philosophical Transactions of the Royal Society B-Biological Sciences*, *367*(1586), doi:10.1098/rstb.2011.0253.
Meirink, J. F., P. Bergamaschi, and M. C. Krol (2008), Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, *Atmospheric Chemistry and Physics*, *8*(21), 6341-6353, doi:10.5194/acp-8-6341-2008
Meng, Z., and F. Zhang (2008), Tests of an Ensemble Kalman Filter for Mesoscale and Regional-Scale Data Assimilation. Part IV: Comparison with 3DVAR in a Month-Long Experiment, *Monthly Weather Review*, *136*(10), 3671-3682, doi:10.1175/2008mwr2270.1.
Meng, Z. Y., and F. Q. Zhang (2011), Limited-Area Ensemble-Based Data Assimilation, *Monthly*

Weather Review, *139*(7), 2025-2045, doi:10.1175/2011mwr3418.1.
Michalak, A. M. (2008), Technical Note: Adapting a fixed-lag Kalman smoother to a geostatistical atmospheric inversion framework, *Atmospheric Chemistry and Physics*, *8*(22), 6789-6799, doi:10.5194/acp-8-6789-2008
Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *Journal of Geophysical Research-Atmospheres*, *109*(D14), doi: 10.1029/2003jd004422.
Michalak, A. M., A. Hirsch, L. Bruhwiler, K. R. Gurney, W. Peters, and P. P. Tans (2005), Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, *Journal of Geophysical Research-Atmospheres*, *110*(D24), doi:10.1029/2005jd005970.
Michel, Y., T. Auligne, and T. Montmerle (2011), Heterogeneous Convective-Scale Background Error Covariances with the Inclusion of Hydrometeor Variables, *Monthly Weather Review*, *139*(9), 2994-3015, doi:10.1175/2011mwr3632.1.
Miller, C. E., et al. (2007), Precision requirements for space-based X-CO2 data, *Journal of Geophysical Research-Atmospheres*, *112*(D10), doi: 10.1029/2006jd007659.
Miyazaki, K., T. Machida, P. K. Patra, T. Iwasaki, Y. Sawa, H. Matsueda, and T. Nakazawa (2009), Formation mechanisms of latitudinal CO2 gradients in the upper troposphere over the subtropics and tropics, *Journal of Geophysical Research-Atmospheres*, *114*, 17, doi:10.1029/2008jd010545.
Miyazaki, K., T. Maki, P. Patra, and T. Nakazawa (2011), Assessing the impact of satellite, aircraft, and surface observations on CO(2) flux estimation using an ensemble-based 4-D data assimilation system, *Journal of Geophysical Research-Atmospheres*, *116*, doi: 10.1029/2010jd015366.
Miyoshi, T. (2011), The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter, *Monthly Weather Review*, *139*(5), 1519-1535, doi:10.1175/2010mwr3570.1.
Miyoshi, T., Y. Sato, and T. Kadowaki (2010), Ensemble Kalman Filter and 4D-Var Intercomparison with the Japanese Operational Global Analysis and Prediction System, *Monthly Weather Review*, *138*(7), 2846-2866, doi:10.1175/2010mwr3209.1.
Montmerle, T., and L. Berre (2010), Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale, *Quarterly Journal of the Royal Meteorological Society*, *136*(651), 1408-1420, doi:10.1002/qj.655.
Moorcroft, P. R. (2006), How close are we to a predictive science of the biosphere?, *Trends in Ecology & Evolution*, *21*(7), 400-407, doi:10.1016/j.tree.2006.04.009.
Mueller, K. L., S. M. Gourdji, and A. M. Michalak (2008), Global monthly averaged CO(2) fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements, *Journal of Geophysical Research-Atmospheres*, *113*(D21), 15, doi: 10.1029/2007jd009734.
Mueller, K. L., V. Yadav, P. S. Curtis, C. Vogel, and A. M. Michalak (2010), Attributing the variability of eddy-covariance CO2 flux measurements across temporal scales using geostatistical regression for a mixed northern hardwood forest, *Global Biogeochemical Cycles*, *24*, doi:10.1029/2009gb003642.
NASA (2008). Active Sensing of CO2 Emissions over Nights, Days, and Seasons (ASCENDS) Mission (http://cce.nasa.gov/ascends/12-30-08%20ASCENDS_Workshop_Report%20clean.pdf). 78. Ann Arbor, Michigan.
NRC (2007) Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond. 11820, National Research Council Committee on Earth Science and Applications from Space: A Community Assessment and Strategy for the Future. Washington D. C.
Nassar, R., et al. (2011), Inverse modeling of CO2 sources and sinks using satellite observations of CO2 from TES and surface flask measurements, *Atmospheric Chemistry and Physics*, *11*(12), doi:10.5194/acp-11-6029-2011.
Nehrkorn, T., J. Eluszkiewicz, S. C. Wofsy, J. C. Lin, C. Gerbig, M. Longo, and S. Freitas (2010), Coupled weather research and forecasting-stochastic time-inverted lagrangian transport (WRF-STILT) model, *Meteorology and Atmospheric Physics*, *107*(1-2), 51-64, doi:10.1007/s00703-010-

0068-x.
Nerger, L., W. Hiller, and J. Schroter (2005), A comparison of error subspace Kalman filters, *Tellus Series a-Dynamic Meteorology and Oceanography*, *57*(5), doi:10.1111/j.1600-0870.2005.00141.x.
Nevison, C. D., N. M. Mahowald, S. C. Doney, I. D. Lima, G. R. Van der Werf, J. T. Randerson, D. F. Baker, P. Kasibhatla, and G. A. McKinley (2008), Contribution of ocean, fossil fuel, land biosphere, and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric CO(2), *Journal of Geophysical Research-Biogeosciences*, *113*(G1), doi:10.1029/2007jg000408.
Nichols, D. (2010), Mathematical Concepts of Data Assimilation, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 13-40, Springer-Verlag Berlin.
Niwa, Y., et al. (2011), Three-dimensional variations of atmospheric CO2: aircraft measurements and multi-transport model simulations, *Atmospheric Chemistry and Physics*, *11*(24), 13359-13375, doi:10.5194/acp-11-13359-2011.
Niwa, Y., T. Machida, Y. Sawa, H. Matsueda, T. J. Schuck, C. A. M. Brenninkmeijer, R. Imasu, and M. Satoh (2012), Imposing strong constraints on tropical terrestrial CO2 fluxes using passenger aircraft based measurements, *Journal of Geophysical Research-Atmospheres*, *117*, doi:10.1029/2012jd017474.
Nocedal, J and S. J. Wright (2006), Numerical Optimization, Springer Ser. Oper. Res., Springer-Verlag, Berlin. pp 224-229
O'Dell, C. W., et al. (2012), The ACOS CO2 retrieval algorithm - Part 1: Description and validation against synthetic observations, *Atmospheric Measurement Techniques*, *5*(1), 99-121, doi:10.5194/amt-5-99-2012.
O'Kane, T. J., and J. S. Frederiksen (2008), Comparison of Statistical Dynamical, Square Root and Ensemble Kalman Filters, *Entropy*, *10*(4), 684-721, doi:10.3390/e10040684.
Oda, T. & S. Maksyutov (2011), A very high-resolution (1km x 1km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of night-time lights, *Atmospheric Chemistry and Physics*, 11, 543-556, doi:10.5194/acp-11-543-2011
Ogata, A and R. B. Banks (1961) A Solution of the Differential Equation of Longitudinal in Porous Media, U.S. Geological Survey Professional Paper 411-A.
Oke, P. R., P. Sakov, and S. P. Corney (2007), Impacts of localisation in the EnKF and EnOI: experiments with a small model, *Ocean Dynamics*, *57*(1), 32-45, doi:10.1007/s10236-006-0088-8.
Olivier, J.G.J. and J. Berdowski (2001), Global Emissions Sources and Sinks in *The Climate System*, edited by J. Berdowski, R. Guicherit, and B. Heij, A. A. Balkema Publishers/Swets and Zeitlinger Publishers, Lisse, The Netherlands, pp 33-78.
Olsen, S. C., and J. T. Randerson (2004), Differences between surface and column atmospheric CO2 and implications for carbon cycle research, *Journal of Geophysical Research-Atmospheres*, *109*(D2), doi:10.1029/2003jd003968.
Olson, D. M., et al. (2001), Terrestrial ecoregions of the worlds: A new map of life on Earth, *Bioscience*, *51*(11), 933-938, doi:10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2.
Omernik, J. M. (2004), Perspectives on the nature and definition of ecological regions, *Environmental Management*, *34*, S27-S38, doi:10.1007/s00267-003-5197-2.
Oshchepkov, S., et al. (2012), Effects of atmospheric light scattering on spectroscopic observations of greenhouse gases from space: Validation of PPDF-based CO2 retrievals from GOSAT, *Journal of Geophysical Research-Atmospheres*, *117*, doi:10.1029/2012jd017505.
Ott, E., B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke (2004), A local ensemble Kalman filter for atmospheric data assimilation, *Tellus Series a-Dynamic Meteorology and Oceanography*, *56*(5), 415-428, doi:10.1111/j.1600-0870.2004.00076.x.
Pacala, S. W., et al. (2001), Consistent land- and atmosphere-based US carbon sink estimates, *Science*, *292*(5525), 2316-2320, doi:10.1126/science.1057320.
Pagano, T. S., M. T. Chahine, and E. T. Olsen (2011), Seven years of observations of mid-tropospheric CO2 from the Atmospheric Infrared Sounder, *Acta Astronautica*, *69*(7-8), 355-359, doi:10.1016/j.actaastro.2011.05.016.

Palmer, P. I., L. Feng, and H. Boesch (2011), Spatial resolution of tropical terrestrial CO2 fluxes inferred using space-borne column CO2 sampled in different earth orbits: the role of spatial error correlations, *Atmospheric Measurement Techniques*, *4*(9), doi:10.5194/amt-4-1995-2011.

Pan, Y. D., et al. (2011), A Large and Persistent Carbon Sink in the World's Forests, *Science*, *333*(6045), 988-993, doi:10.1126/science.1201609.

Pannekoucke, O., L. Berre, and G. Desroziers (2007), Filtering properties of wavelets for local background-error correlations, *Quarterly Journal of the Royal Meteorological Society*, *133*(623), 363-379, doi:10.1002/qj.33.

Pannekoucke, O., L. Berre, and G. Desroziers (2008), Background-error correlation length-scale estimates and their sampling statistics, *Quarterly Journal of the Royal Meteorological Society*, *134*(631), 497-508, doi:10.1002/qj.212.

Parazoo, N. C., A. S. Denning, S. R. Kawa, K. D. Corbin, R. S. Lokupitiya, and I. T. Baker (2008), Mechanisms for synoptic variations of atmospheric CO2 in North America, South America and Europe, *Atmospheric Chemistry and Physics*, *8*(23)., doi:10.5194/acp-8-7239-2008

Parazoo, N. C., A. S. Denning, S. R. Kawa, S. Pawson, and R. Lokupitiya (2012), CO2 flux estimation errors associated with moist atmospheric processes, *Atmospheric Chemistry and Physics*, *12*(14), 6405-6416, doi:10.5194/acp-12-6405-2012.

Parrish, D. F., and J. C. Derber (1992), The National Meteorological Centers Spectral Statistical Interpolation Analysis System, *Monthly Weather Review*, *120*(8), 1747-1763, doi:10.1175/1520-0493(1992)120<1747:tnmcss>2.0.co;2.

Patra, P. K., Y. Niwa, T. J. Schuck, C. A. M. Brenninkmeijer, T. Machida, H. Matsueda, and Y. Sawa (2011), Carbon balance of South Asia constrained by passenger aircraft CO2 measurements, *Atmospheric Chemistry and Physics*, *11*(9), doi:10.5194/acp-11-4163-2011.

Patra, P. K., J. G. Canadell, and S. Lal (2012), The rapidly changing greenhouse gas budget of Asia, *Eos Trans. AGU*, 93(25), 237, doi:10.1029/2012EO250006

Pena, M., Z. Toth, and M. Z. Wei (2010), Controlling Noise in Ensemble Data Assimilation Schemes, *Monthly Weather Review*, *138*(5), 1502-1512, doi:10.1175/2009mwr2854.1.

Peters, W., et al. (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, *Proceedings of the National Academy of Sciences of the United States of America*, *104*(48), 18925-18930, doi:10.1073/pnas.0708986104.

Peters, W., et al. (2010), Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biology*, *16*(4), 1317-1337, doi:10.1111/j.1365-2486.2009.02078.x.

Peters, W., J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans (2005), An ensemble data assimilation system to estimate CO2 surface fluxes from atmospheric trace gas observations, *Journal of Geophysical Research-Atmospheres*, *110*(D24), doi: 10.1029/2005jd006157.

Petron, G., C. Granier, B. Khattatov, V. Yudin, J. F. Lamarque, L. Emmons, J. Gille, and D. P. Edwards (2004), Monthly CO surface sources inventory based on the 2000-2001 MOPITT satellite data, *Geophysical Research Letters*, *31*(21), doi:10.1029/2004gl020560.

Peylin, P., D. Baker, J. Sarmiento, P. Ciais, and P. Bousquet (2002), Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO2 data, *Journal of Geophysical Research-Atmospheres*, *107*(D19), doi:10.1029/2001jd000857.

Peylin, P., P. Bousquet, C. Le Quere, S. Sitch, P. Friedlingstein, G. McKinley, N. Gruber, P. Rayner, and P. Ciais (2005), Multiple constraints on regional CO2 flux variations over land and oceans, *Global Biogeochemical Cycles*, *19*(1), 24, doi:10.1029/2003gb002214.

Pham, D. T. (2001), Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Monthly Weather Review*, *129*(5), 1194-1207, doi: 10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2

Phillips, O. L., N. Higuchi, S. Vieira, T. R. Baker, K.-J. Chao, and S. L. Lewis (2009), Changes in Amazonian Forest Biomass, Dynamics, and Composition, 1980-2002, *Amazonia and Global Change*,

186, doi:10.1029/2008gm000739.
Piao, S. L., P. Ciais, P. Friedlingstein, N. de Noblet-Ducoudre, P. Cadule, N. Viovy, and T. Wang (2009), Spatiotemporal patterns of terrestrial carbon cycle during the 20th century, *Global Biogeochemical Cycles*, *23*, 16, doi:10.1029/2008gb003339.
Pickett-Heaps, C. A., et al. (2011), Atmospheric CO2 inversion validation using vertical profile measurements: Analysis of four independent inversion models, *Journal of Geophysical Research-Atmospheres*, *116*, doi:10.1029/2010jd014887.
Poli, P., J. Joiner, and E. R. Kursinski (2002), 1DVAR analysis of temperature and humidity using GPS radio occultation refractivity data, *Journal of Geophysical Research-Atmospheres*, *107*(D20), doi:10.1029/2001jd000935.
Potter, C., S. Klooster, A. Huete, and V. Genovese (2007), Terrestrial carbon sinks for the United States predicted from MODIS satellite data and ecosystem modeling, *Earth Interactions*, *11*, doi:10.1175/ei228.1.
Potter, C., S. Klooster, R. Myneni, V. Genovese, P. N. Tan, and V. Kumar (2003), Continental-scale comparisons of terrestrial carbon sinks estimated from satellite data and ecosystem modeling 1982-1998, *Global and Planetary Change*, *39*(3-4), 201-213, doi:10.1016/j.gloplacha.2003.07.001.
Potter, C. S., J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney, and S. A. Klooster (1993), Terrestrial Ecosystem Production – A Process Model-Based on Global Satellite and Surface Data, *Global Biogeochemical Cycles*, *7*(4), doi:10.1029/93gb02725.
Rabier, F. (2005), Overview of global data assimilation developments in numerical weather-prediction centres, *Quarterly Journal of the Royal Meteorological Society*, *131*(613), doi:10.1256/qj.05.129.
Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biology*, *15*(10), 2462-2484, doi:10.1111/j.1365-2486.2009.01912.x.
Randerson, J. T., M. V. Thompson, T. J. Conway, I. Y. Fung, and C. B. Field (1997), The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochemical Cycles*, *11*(4), 535-560, doi:10.1029/97gb02268.
Raupach, M. R., and J. G. Canadell (2010), Carbon and the Anthropocene, *Current Opinion in Environmental Sustainability*, *2*(4), doi:10.1016/j.cosust.2010.04.003.
Raupach, M. R., J. G. Canadell, P. Ciais, P. Friedlingstein, P. J. Rayner, and C. M. Trudinger (2011), The relationship between peak warming and cumulative CO(2) emissions, and its use to quantify vulnerabilities in the carbon-climate-human system, *Tellus Series B-Chemical and Physical Meteorology*, *63*(2), 145-164, doi:10.1111/j.1600-0889.2010.00521.x.
Raupach, M. R., and J. J. Finnigan (1995), Scale Issues in Boundary-Layer Meteorology – Surface-Energy Balances in Heterogeneous Terrain, *Hydrological Processes*, *9*(5-6), doi:10.1002/hyp.3360090509.
Raynaud, L., L. Berre, and G. Desroziers (2009), Objective filtering of ensemble-based background-error variances, *Quarterly Journal of the Royal Meteorological Society*, *135*(642), 1177-1199, doi:10.1002/qj.438.
Rayner, P. J. (2010), The current state of carbon-cycle data assimilation, *Current Opinion in Environmental Sustainability*, *2*(4), 289-296, doi:10.1016/j.cosust.2010.05.005.
Rayner, P. J., E. Koffi, M. Scholze, T. Kaminski, and J. L. Dufresne (2011), Constraining predictions of the carbon cycle using data, *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, *369*(1943), 1955-1966, doi:10.1098/rsta.2010.0378.
Rayner, P. J., R. M. Law, C. E. Allison, R. J. Francey, C. M. Trudinger, and C. Pickett-Heaps (2008), Interannual variability of the global carbon cycle (1992-2005) inferred by inversion of atmospheric CO(2) and delta(13)CO(2) measurements, *Global Biogeochemical Cycles*, *22*(3), doi:10.1029/2007gb003068.
Rayner, P. J., and D. M. O'Brien (2001), The utility of remotely sensed CO2 concentration data in surface source inversions (vol 28, pg 175, 2001), *Geophysical Research Letters*, *28*(12), 2429-2429, doi:10.1029/2001gl013115.
Rayner, P. J., M. Scholze, W. Knorr, T. Kaminski, R. Giering, and H. Widmann (2005), Two decades of

terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), *Global Biogeochemical Cycles*, *19*(2), doi: 10.1029/2004gb002254.
Reichle, R. H. (2008), Data assimilation methods in the Earth sciences, *Advances in Water Resources*, *31*(11), doi:10.1016/j.advwatres.2008.01.001.
Revelle, R., and H. E. Suess (1957), Carbon Dioxide Exchange between Atmosphere and Ocean and the Question of an increase of Atmospheric CO2 during the past decades, *Tellus*, *9*(1). 18-27, doi: 10.1111/j.2153-3490.1957.tb01849.x
Richardson, A. D., et al. (2006), A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, *Agricultural and Forest Meteorology*, *136*(1-2), 1-18, doi:10.1016/j.agrformet.2006.01.007.
Richardson, A. D., et al. (2008), Statistical properties of random CO2 flux measurement uncertainty inferred from model residuals, *Agricultural and Forest Meteorology*, *148*(1), 38-50, doi:10.1016/j.agrformet.2007.09.001.
Rivier, L., et al. (2010), European CO2 fluxes from atmospheric inversions using regional and global transport models, *Climatic Change*, *103*(1-2), doi:10.1007/s10584-010-9908-4.
Rodell, M., et al. (2004), The global land data assimilation system, *Bulletin of the American Meteorological Society*, *85*(3), 381-+, doi:10.1175/bams-85-3-381.
Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann (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, doi:10.5194/acp-3-1919-2003
Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003b), Time-dependent atmospheric CO2 inversions based on interannually varying tracer transport, *Tellus Series B-Chemical and Physical Meteorology*, *55*(2), doi:10.1034/j.1600-0889.2003.00033.x.
Rodenbeck, C. (2005), Estimating CO2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Max-Planck-Institut fur Biogeochemie: Technical Paper 6.
Rodgers, C.D (2000) *Inverse methods for atmospheric sounding: theory and practice.* World Scientific Publ., Singapore, 238 pp.
Roebber, P. J., M. Westendorf, and G. R. Meadows (2010), Innovative Weather: A New Strategy for Student, University, and Community Relationships, *Bulletin of the American Meteorological Society*, *91*(7), 877-888, doi:10.1175/2010bams2854.1.
Rood, R. B. (2005), Assimilation of stratospheric meteorological and constituent observations: A Review. *SPARC Newsletter no. 25*, Available at: http://www.atmosp.physics.utoronto.ca/SPARC/News25/Assimilation%20Rood.html
Roy, T. et al. (2011), Regional Impacts of Climate Change and Atmospheric CO2 on Future Ocean Carbon Uptake: A Multimodel Linear Feedback Analysis, *Journal of Climate*, *24*(9), 2300–2318, doi: 10.1175/2010JCLI3787.1
Ruimy, A., L. Kergoat, C. B. Field, and B. Saugier (1996), The use of CO2 flux measurements in models of the global terrestrial carbon budget, *Global Change Biology*, *2*(3), 287-296, doi:10.1111/j.1365-2486.1996.tb00080.x.
Runkel, R. L. (1996), Solution of the advection-dispersion equation: Continuous load of finite duration, *Journal of Environmental Engineering-Asce*, *122*(9), doi:10.1061/(asce)0733-9372(1996)122:9(830).
Sabine, C. L., et al. (2004), The oceanic sink for anthropogenic CO(2), *Science*, *305*(5682), 367-371, doi:10.1126/science.1097403.
Sacher, W., and P. Bartello (2008), Sampling errors in ensemble Kalman filtering. Part I: Theory, *Monthly Weather Review*, *136*(8), 3035-3049, doi:10.1175/2007mwr2323.1.
Saito, R., S. Houweling, P. K. Patra, D. Belikov, R. Lokupitiya, Y. Niwa, F. Chevallier, T. Saeki, and S. Maksyutov (2011), TransCom satellite intercomparison experiment: Construction of a bias corrected atmospheric CO2 climatology, *Journal of Geophysical Research-Atmospheres*, *116*, 13, doi:10.1029/2011jd016033.
Saleska, S. R., et al. (2003), Carbon in amazon forests: Unexpected seasonal fluxes and disturbance-

induced losses, *Science*, *302*(5650), doi:10.1126/science.1091165.
Santaren, D., P. Peylin, N. Viovy, and P. Ciais (2007), Optimizing a process-based ecosystem model with eddy-covariance flux measurements: A pine forest in southern France, *Global Biogeochemical Cycles*, *21*(2), doi:10.1029/2006gb002834.
Sarmiento, J. L., M. Gloor, N. Gruber, C. Beaulieu, A. R. Jacobson, S. E. M. Fletcher, S. Pacala, and K. Rodgers (2010), Trends and regional distributions of land and ocean carbon sinks, *Biogeosciences*, *7*(8), 2351-2367, doi:10.5194/bg-7-2351-2010.
Sarrat, C., et al. (2007), Atmospheric CO2 modeling at the regional scale: an intercomparison of 5 meso-scale atmospheric models, *Biogeosciences*, *4*(6), doi:10.5194/bg-4-1115-2007
Sasaki, Y. (1970a), Numerical Variational Analysis Formulated Under Constraints as Determined By Long wave Equations and a Low-pass Filter, *Monthly Weather Review*, *98*(12), doi:10.1175/1520-0493(1970)098<0884:nvafut>2.3.co;2.
Sasaki, Y. (1970b), Numerical Variational Analysis with Weak Constraint and Application to Surface Analysis of Severe Storm Gust, *Monthly Weather Review*, *98*(12), doi:10.1175/1520-0493(1970)098<0899:nvawwc>2.3.co;2.
Sasaki, Y. (1970c), Some Basic Formalisms in Numerical Variational Analysis, *Monthly Weather Review*, *98*(12), doi:10.1175/1520-0493(1970)098<0875:sbfinv>2.3.co;2.
Sawa, Y., T. Machida, and H. Matsueda (2012), Aircraft observation of the seasonal variation in the transport of CO2 in the upper atmosphere, *Journal of Geophysical Research-Atmospheres*, *117*, doi:10.1029/2011jd016933.
Schimel, D. (2007), Carbon cycle conundrums, *Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), 18353-18354, doi:10.1073/pnas.0709331104.
Schneising, O., et al. (2012), Atmospheric greenhouse gases retrieved from SCIAMACHY: comparison to ground-based FTS measurements and model results, *Atmospheric Chemistry and Physics*, *12*(3), 1527-1540, doi:10.5194/acp-12-1527-2012.
Scholes, R. J., P. M. S. Monteiro, C. L. Sabine, and J. G. Canadell (2009), Systematic long-term observations of the global carbon cycle, *Trends in Ecology & Evolution*, *24*(8), 427-430, doi:10.1016/j.tree.2009.03.006.
Scholze, M., T. Kaminski, P. Rayner, W. Knorr, and R. Giering (2007), Propagating uncertainty through prognostic carbon cycle data assimilation system simulations, *Journal of Geophysical Research-Atmospheres*, *112*(D17), doi: 10.1029/2007jd008642.
Schuh, A. E., A. S. Denning, K. D. Corbin, I. T. Baker, M. Uliasz, N. Parazoo, A. E. Andrews, and D. E. J. Worthy (2010), A regional high-resolution carbon flux inversion of North America for 2004, *Biogeosciences*, *7*(5), doi:10.5194/bg-7-1625-2010.
Schutgens, N. A. J., T. Miyoshi, T. Takemura, and T. Nakajima (2010), Sensitivity tests for an ensemble Kalman filter for aerosol assimilation, *Atmospheric Chemistry and Physics*, *10*(14), doi:10.5194/acp-10-6583-2010.
Schwalm, C. R., et al. (2010), A model-data intercomparison of CO2 exchange across North America: Results from the North American Carbon Program site synthesis, *Journal of Geophysical Research-Biogeosciences*, *115*, doi:10.1029/2009jg001229.
Schwarz, G. (1978), Estimating Dimension of a Model, *Annals of Statistics*, *6*(2), doi:10.1214/aos/1176344136.
Shutyaev, V. P., F. X. Le Dimet, and I. Y. Gejadze (2009), A posteriori error covariances in variational data assimilation, *Russian Journal of Numerical Analysis and Mathematical Modelling*, *24*(2), 161-169, doi:10.1515/rjnamm.2009.011.
Singh, K., M. Jardak, A. Sandu, K. Bowman, M. Lee, and D. Jones (2011), Construction of non-diagonal background error covariance matrices for global chemical data assimilation, *Geoscientific Model Development*, *4*(2), 299-316, doi:10.5194/gmd-4-299-2011.
Siroka, M., C. Fischer, V. Casse, R. Brozkova, and J. F. Geleyn (2003), The definition of mesoscale selective forecast error covariances for a limited area variational analysis, *Meteorology and Atmospheric Physics*, *82*(1-4), 227-244, doi:10.1007/s00703-001-0588-5.

Sitch, S., et al. (2008), Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), *Global Change Biology*, *14*(9), doi:10.1111/j.1365-2486.2008.01626.x.

Skamarock, W. C., et. al. (2005).A description of the advanced research WRF version 2. Technical Note 468+STR, MMM Division, NCAR, Boulder, CO. Available from: http://www.mmm.ucar.edu/wrf/users/docs/arw_v2.pdf

Snyder, C., T. Bengtsson, P. Bickel, and J. Anderson (2008), Obstacles to High-Dimensional Particle Filtering, *Monthly Weather Review*, *136*(12), 4629-4640, doi:10.1175/2008mwr2529.1.

State of the Carbon Cycle Report (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., L. Dilling, G.P. Zimmerman, D.M. Fairman, R.A. Houghton, G. Marland, A.Z. Rose, & T.J. Wilbanks (eds.)). National Oceanic and Atmospheric Administration, National Climatic Data Center, Asheville, NC, USA, 242 pp.

Stephens, B. B., et al. (2007), Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO(2), *Science*, *316*(5832), 1732-1735, doi:10.1126/science.1137004.

Storto, A., and R. Randriamampianina (2010), Ensemble variational assimilation for the representation of background error covariances in a high-latitude regional model, *Journal of Geophysical Research-Atmospheres*, *115*, 21, doi:10.1029/2009jd013111.

Swinbank, R (2010), Numerical Weather Prediction, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 381-407, Springer-Verlag Berlin.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, D. J. Patil, B. R. Hunt, E. Kalnay, E. Ott, and J. A. Yorke (2005), Assessing a local ensemble Kalman filter: perfect model experiments with the National Centers for Environmental Prediction global model, *Tellus Series a-Dynamic Meteorology and Oceanography*, *57*(4), 528-545, doi:10.1111/j.1600-0870.2005.00136.x.

Takagi, H., et al. (2011), On the Benefit of GOSAT Observations to the Estimation of Regional CO2 Fluxes, *Sola*, *7*, doi:10.2151/sola.2011-041.

Takahashi, T., et al. (2002), Global sea-air CO2 flux based on climatological surface ocean pCO(2), and seasonal biological and temperature effects, *Deep-Sea Research Part Ii-Topical Studies in Oceanography*, *49*(9-10), doi:10.1016/s0967-0645(02)00003-6.

Talagrand, O., and P. Courtier (1987), Variational Assimilation of Meteorological Observations with the Adjoint Vorticity Equation. 1. Theory, *Quarterly Journal of the Royal Meteorological Society*, *113*(478), doi:10.1256/smsqj.47811.

Talagrand,O. (2010), Variational Assimilation, in *Data Assimilation, Making Sense of Observations*, edited by W. Lahoz, B. Khattatov, R. Menard, pp 41-67, Springer-Verlag Berlin.

Tans, P. P. (1996), Carbon cycle (group report), in Summary Report 1994–1995, vol. 23, edited by D. J. Hoffman, J. Peterson, and R. M. Rosson, pp. 29–49, U. S. Dep. of Commer, Boulder, Colorado.

Tans, P., and T. J. Conway (2005), Monthly Atmospheric CO2 Mixing Ratios from the NOAA CMDL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2002., In *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Tarantola, A. (2005). Inverse Problem Theory and Methods for Model Parameter Estimation, *Society for Industrial and Applied Mathematics*, 358 pp.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research-Atmospheres*, *106*(D7), doi:10.1029/2000jd900719.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker (2003), Ensemble square root filters, *Monthly Weather Review*, *131*(7), 1485-1490, doi: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2

Tiwari, Y. K., M. Gloor, R. J. Engelen, F. Chevallier, C. Roedenbeck, S. Koerner, P. Peylin, B. H. Braswell, and M. Heimann (2006), Comparing CO2 retrieved from Atmospheric Infrared Sounder with model predictions: Implications for constraining surface fluxes and lower-to-upper troposphere

transport, *Journal of Geophysical Research-Atmospheres*, *111*(D17), doi:10.1029/2005jd006681.
Todling, R. (1999). Estimation Theory and foundations of atmospheric data assimilation, NASA/DAO Office notes, 183 pp. http://gmao.gsfc.nasa.gov/pubs/docs/Todling180.pdf
Tremolet, Y. (2006), Accounting for an imperfect model in 4D-Var, *Quarterly Journal of the Royal Meteorological Society*, *132*(621), doi:10.1256/qj.05.224.
Tremolet, Y. (2007), Model-error estimation in 4D-Var, *Quarterly Journal of the Royal Meteorological Society*, *133*(626), doi:10.1002/qj.94.
Turner, D. P., M. Gockede, B. E. Law, W. D. Ritts, W. B. Cohen, Z. Yang, T. Hudiburg, R. Kennedy, and M. Duane (2011), Multiple constraint analysis of regional land-surface carbon flux, *Tellus Series B-Chemical and Physical Meteorology*, *63*(2), 207-221, doi:10.1111/j.1600-0889.2011.00525.x.
Tyndall, J. (1861), On the Absorption and Radiation of Heat by Gases and Vapours, and on the Physical Connection of Radiation, Absorption, and Conduction, *Philosophical Magazine* ser. 4, 22, 169–94, 273–85.
U.S. Carbon Cycle Science Plan (2011), A U.S Carbon Cycle Science Plan, A Report of the Carbon Cycle Science Steering Group and Subcommittee, Anna Michalak, Rob Jackson, Gregg Marland, Chris Sabine, Co-Chairs, Available at http://www.carboncyclescience.gov/carbonplanning.php, Accessed on August 11, 2012
U. S. Climate Change Science Program (2009), Our Changing Planet, The US climate change science program for fiscal year 2009, Available at: http://www.usgcrp.gov/usgcrp/Library/ocp2009/default.htm, Accessed on August 11, 2012
Uzunoglu, B., S. J. Fletcher, M. Zupanski, and I. M. Navon (2007), Adaptive ensemble reduction and inflation, *Quarterly Journal of the Royal Meteorological Society*, *133*(626), 1281-1294, doi:10.1002/qj.96.
van der Molen, M. K., and A. J. Dolman (2007), Regional carbon fluxes and the effect of topography on the variability of atmospheric CO2, *Journal of Geophysical Research-Atmospheres*, *112*(D1), 16, doi:10.1029/2006jd007649.
van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmospheric Chemistry and Physics*, *6*, 3423-3441, doi:10.5194/acp-6-3423-2006
van Leeuwen, P. J. (2010), Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Quarterly Journal of the Royal Meteorological Society*, *136*(653), doi:10.1002/qj.699.
vanLeeuwen, P. J., and G. Evensen (1996), Data assimilation and inverse methods in terms of a probabilistic formulation, *Monthly Weather Review*, *124*(12), doi:10.1175/1520-0493(1996)124<2898:daaimi>2.0.co;2.
Varella, H., L. Berre, and G. Desroziers (2011), Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model, *Quarterly Journal of the Royal Meteorological Society*, *137*(658), 1369-1379, doi:10.1002/qj.845.
Vukicevic, T., E. Kalnay, and T. Vonder Haar (2004), The need for a national data assimilation education program, *Bulletin of the American Meteorological Society*, *85*(1), doi:10.1175/bams-85-1-48.
Wang, B., J. J. Liu, S. D. Wang, W. Cheng, J. A. Liu, C. S. Liu, Q. N. Xiao, and Y. H. Kuo (2010), An economical approach to four-dimensional variational data assimilation, *Advances in Atmospheric Sciences*, *27*(4), 715-727, doi:10.1007/s00376-009-9122-3.
Wang, X., T. A. Hamill, J. S. Whitaker, and C. H. Bishop (2007), A comparison of hybrid ensemble transform Kalman filter-optimum interpolation and ensemble square root filter analysis schemes, *Monthly Weather Review*, *135*(3), doi:10.1175/mwr3307.1.
Wang, X. G., and C. H. Bishop (2003), A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes, *Journal of the Atmospheric Sciences*, *60*(9), doi:10.1175/1520-0469(2003)060<1140:acobae>2.0.co;2.
Wang, Y. P., C. M. Trudinger, and I. G. Enting (2009), A review of applications of model-data fusion to studies of terrestrial carbon fluxes at different scales, *Agricultural and Forest Meteorology*, *149*(11), 1829-1842, doi:10.1016/j.agrformet.2009.07.009.

Wanninkhof, R., W. E. Asher, D. T. Ho, C. Sweeney, and W. R. McGillis (2009), Advances in Quantifying Air-Sea Gas Exchange and Environmental Forcing, in *Annual Review of Marine Science*, edited, pp. 213-244, Annual Reviews, Palo Alto, doi:10.1146/annurev.marine.010908.163742.

Whitaker, J. S. and G. P. Compo (2002), An ensemble Kalman smoother for reanalysis, *Symposium on Observations, Data Assimilation, and Probabilistic Prediction*, Available at: https://ams.confex.com/ams/pdfpapers/28864.pdf

Whitaker, J. S., G. P. Compo, and J. N. Thepaut (2009), A Comparison of Variational and Ensemble-Based Data Assimilation Systems for Reanalysis of Sparse Observations, *Monthly Weather Review*, *137*(6), 1991-1999, doi:10.1175/2008mwr2781.1.

Whitaker, J. S., and T. M. Hamill (2002), Ensemble data assimilation without perturbed observations, *Monthly Weather Review*, *130*(7), 1913-1924, doi:10.1175/1520-0493(2002)130<1913:edawpo>2.0.co;2.

Whitaker, J. S., and A. F. Loughe (1998), The relationship between ensemble spread and ensemble mean skill, *Monthly Weather Review*, *126*(12), 3292-3302, doi:10.1175/1520-0493(1998)126<3292:trbesa>2.0.co;2.

Williams, M., et al. (2009), Improving land surface models with FLUXNET data, *Biogeosciences*, *6*(7), 1341-1359, doi:10.5194/bg-6-1341-2009

Willmott, C. J., and K. Matsuura (2005), Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Research*, *30*(1), doi:10.3354/cr030079.

Wofsy, S. C., M. L. Goulden, J. W. Munger, S. M. Fan, P. S. Bakwin, B. C. Daube, S. L. Bassow, and F. A. Bazzaz (1993), Net Exchange of CO2 in a Mid-latitude Forest, *Science*, *260*(5112), doi:10.1126/science.260.5112.1314.

Wu, L., V. Mallet, M. Bocquet, and B. Sportisse (2008), A comparison study of data assimilation algorithms for ozone forecasts, *Journal of Geophysical Research-Atmospheres*, *113*(D20), doi:10.1029/2008jd009991.

Wulder, M. A., J. C. White, R. A. Fournier, J. E. Luther, and S. Magnussen (2008), Spatially explicit large area biomass estimation: Three approaches using forest inventory and remotely sensed imagery in a GIS, *Sensors*, *8*(1), 529-560, doi:10.3390/s8010529.

Wunch, D., G. C. Toon, J.-F. L. Blavier, R. A. Washenfelder, J. Notholt, B. J. Connor, D. W. T. Griffith, V. Sherlock, and P. O. Wennberg (2011a), The Total Carbon Column Observing Network, *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences*, *369*(1943), 2087-2112, doi:10.1098/rsta.2010.0240.

Wunch, D., et al. (2011b), A method for evaluating bias in global measurements of CO2 total columns from space, *Atmospheric Chemistry and Physics*, *11*(23), doi:10.5194/acp-11-12317-2011.

Xiao, J., J. Chen, K. J. Davis, and M. Reichstein (2012), Advances in upscaling of eddy covariance measurements of carbon and water fluxes, *Journal of Geophysical Research – Biogeosciences*, 117, G00J01, doi:10.1029/2011JG001889.

Xu, T., L. White, D. F. Hui, and Y. Q. Luo (2006), Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction, *Global Biogeochemical Cycles*, *20*(2), 15, doi:10.1029/2005gb002468.

Xueref-Remy, I., P. Bousquet, C. Carouge, L. Rivier, and P. Ciais (2011a), Variability and budget of CO2 in Europe: analysis of the CAATER airborne campaigns - Part 2: Comparison of CO2 vertical variability and fluxes between observations and a modeling framework, *Atmospheric Chemistry and Physics*, *11*(12), doi:10.5194/acp-11-5673-2011.

Xueref-Remy, I., C. Messager, D. Filippi, M. Pastel, P. Nedelec, M. Ramonet, J. D. Paris, and P. Ciais (2011b), Variability and budget of CO2 in Europe: analysis of the CAATER airborne campaigns - Part 1: Observed variability, *Atmospheric Chemistry and Physics*, *11*(12), doi:10.5194/acp-11-5655-2011.

Yadav, V., K. L. Mueller, D. Dragoni, and A. M. Michalak (2010a), A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America, *Biogeosciences*,

7(9), doi:10.5194/bg-7-2655-2010.
Yadav, V., C. J. Antonelli, S. M. Gourdji, K. L. Mueller, A. Chatterjee, A. M. Michalak (2010b), Design Framework for a Real-Time Large-Scale, Parallel, Intelligent, CO_2 Data Assimilation System, Computational Discovery and Cyber-Infrastructure at University of Michigan, Ann Arbor, Accessible at- http://orci.research.umich.edu/cidays/cyberinfrastructure-days-2010/poster-session/
Yokota, T., Y. Yoshida, N. Eguchi, Y. Ota, T. Tanaka, H. Watanabe, and S. Maksyutov (2009), Global Concentrations of CO2 and CH4 Retrieved from GOSAT: First Preliminary Results, *Sola*, *5*, doi:10.2151/sola.2009-041.
Yoshida, Y., Y. Ota, N. Eguchi, N. Kikuchi, K. Nobuta, H. Tran, I. Morino, and T. Yokota (2011), Retrieval algorithm for CO2 and CH4 column abundances from short-wavelength infrared spectral observations by the Greenhouse gases observing satellite, *Atmospheric Measurement Techniques*, *4*(4), doi:10.5194/amt-4-717-2011.
Yudin, V. A., G. Petron, J. F. Lamarque, B. V. Khattatov, P. G. Hess, L. V. Lyjak, J. C. Gille, D. P. Edwards, M. N. Deeter, and L. K. Emmons (2004), Assimilation of the 2000-2001 CO MOPITT retrievals with optimized surface emissions, *Geophysical Research Letters*, *31*(20), doi:10.1029/2004gl021037.
Zagar, N., E. Andersson, and M. Fisher (2005), Balanced tropical data assimilation based on a study of equatorial waves in ECMWF short-range forecast errors, *Quarterly Journal of the Royal Meteorological Society*, *131*(607), 987-1011, doi:10.1256/qj.04.54.
Zhang, F. Q., C. Snyder, and J. Z. Sun (2004), Impacts of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter, *Monthly Weather Review*, *132*(5), 1238-1253, doi:10.1175/1520-0493(2004)132<1238:ioieao>2.0.co;2.
Zhang, M., F. Q. Zhang, X. Y. Huang, and X. Zhang (2011), Intercomparison of an Ensemble Kalman Filter with Three- and Four-Dimensional Variational Data Assimilation Methods in a Limited-Area Model over the Month of June 2003, *Monthly Weather Review*, *139*(2), 566-572, doi:10.1175/2010mwr3610.1.
Zhiyong, M., and Z. Fuqing (2008), Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part III: comparison with 3DVAR in a real-data case study, *Monthly Weather Review*, *136*(2), doi:10.1175/2007mwr2106.1.
Zhuravlev, R., B. Khattatov, B. Kiryushov, and S. Maksyutov (2011), Technical Note: A novel approach to estimation of time-variable surface sources and sinks of carbon dioxide using empirical orthogonal functions and the Kalman filter, *Atmospheric Chemistry and Physics*, *11*(20), doi:10.5194/acp-11-10305-2011.
Ziehn, T., J. Kattge, W. Knorr, and M. Scholze (2011a), Improving the predictability of global doi:10.1029/2011gl047182.
Ziehn, T., W. Knorr, and M. Scholze (2011b), Investigating spatial differentiation of model parameters in a carbon cycle data assimilation system, *Global Biogeochemical Cycles*, *25*, doi:10.1029/2010gb003886.
Zupanski, D., A. S. Denning, M. Uliasz, M. Zupanski, A. E. Schuh, P. J. Rayner, W. Peters, and K. D. Corbin (2007a), Carbon flux bias estimation employing maximum likelihood ensemble filter (MLEF), *Journal of Geophysical Research-Atmospheres*, *112*(D17), doi:10.1029/2006jd008371.
Zupanski, D., A. Y. Hou, S. Q. Zhang, M. Zupanski, C. D. Kummerow, and S. H. Cheung (2007b), Applications of information theory in ensemble data assimilation, *Quarterly Journal of the Royal Meteorological Society*, *133*(627), doi:10.1002/qj.123.
Zupanski, M. (2005), Maximum likelihood ensemble filter: Theoretical aspects, *Monthly Weather Review*, *133*(6), 1710-1726, doi:10.1175/mwr2946.1.