Improved estimates of regional-scale land-atmosphere CO_2 exchange using geostatistical atmospheric inverse models

By

Sharon Muzli Gourdji

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Environmental Engineering)
in the University of Michigan
2011

Doctoral committee:

Associate Professor Anna M. Michalak, Chair
Professor Peter Adriaens
Professor Edward A. Parson
Professor Richard B. Rood
Professor A. Scott Denning, Colorado State University

2011

ACKNOWLEDGEMENTS

The work presented in this dissertation represents the culmination of a 5-year long project involving the collaboration and assistance of many of my fellow scientists at the University of Michigan, the National Oceanic & Atmospheric Administration, and AER Inc. Principally, I would like to acknowledge my fellow Ph.D. student Kim Mueller who worked closely with me on the geostatistical inversions throughout this period. Starting out with little background in the field, together we found a way to interpret and make our results meaningful to the larger scientific community.

I thank my advisor, Dr. Anna Michalak, for taking me on as a Ph.D. student, and for providing the funding and environment in which this work took place. Her insight regarding statistical optimization problems, support and concern for her students, and energy levels are unmatched.

I thank my committee members (Drs. Peter Adriaens, Scott Denning, Ted Parson and Ricky Rood) for making the effort to learn about my work and provide valuable direction, especially given that most members were not specifically familiar with the technical aspects of the work.

I thank my research group (PUORG) for numerous sessions of feedback in research group meetings and on manuscript reviews. The synergy between all of our research activities inspired me, and helped me to understand my own work in the broader context of carbon cycle science and spatial statistics. I also thank the members of my group individually for their encouragement, support and friendship throughout this Ph.D.

I thank the funding agencies that financially supported this work. These include grants from the National Oceanic and Atmospheric Administration Earth System

Research Laboratory (RA133R-05-SE-5150, or “Geostatistical analysis of NOAA Climate Monitoring and Diagnostics Laboratory Carbon Dioxide Data for 1997-2001”), and the National Aeronautics and Space Administration (NNX06AE84G, or “Constraining North American Fluxes of Carbon Dioxide and Inferring their Spatiotemporal Covariances through Assimilation of Remote Sensing and Atmospheric Data in a Geostatistical Framework” issued through the ROSES A.6 North American Carbon Program). I also received additional support through a NASA Earth System Science Fellowship.

Finally, last but not least, I thank my parents, Judith and Alfred Gourdji, and Vineeth Veetil for their personal encouragement and love throughout this Ph.D. Thank you also to my grandmother, Rae Mike, who encouraged me to study math in college to damp down my political passions. ☺ I know that you all take pride in my accomplishments.

Portions of this work were adapted from the following published studies or manuscripts in preparation:

- Gourdji, S.M., K.L. Mueller, K. Schaefer, & A.M. Michalak (2008), Global monthly-averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *J. Geophys. Res.*, 113(D21115), doi:10.1029/2007JD009733.

- Gourdji, S.M., A.I. Hirsch, K.L. Mueller, V. Yadav, A.E. Andrews, & A.M. Michalak (2010), Regional-scale geostatistical inverse modeling of North American CO2 fluxes: a synthetic data study, *Atm. Chem. Phys.*, 10, 6151-6167.

- Gourdji, S.M., K.L. Mueller, V. Yadav, D.N. Huntzinger, A.E. Andrews, M. Trudeau, G. Petron, T. Nehrkorn, J. Eluskiewicz, J. Henderson, D. Wen, J. Lin, D. Worthy, W. Munger, M. Fischer, & A.M. Michalak (in prep.), Results from a regional grid-scale North American atmospheric CO_2 inversion for 2004 with a comparison to independent bottom-up flux estimates.

The co-authors of these manuscripts are gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Improved estimates of regional-scale land-atmosphere CO_2 exchange using geostatistical atmospheric inverse models

by

Sharon Muzli Gourdji

Chair: Anna M. Michalak

In order to devise strategies to reduce atmospheric CO_2 concentrations and predict their future trajectories for climate change mitigation and prediction, it is important to accurately quantify and understand the drivers of regional-scale (~500 x 500 km^2) land-atmosphere carbon exchange from biospheric processes and fossil fuel emissions. While CO_2 fluxes at this scale cannot be directly measured, inverse models can potentially provide estimates with reasonable uncertainties by tracing back variability in atmospheric CO_2 measurements to the most likely distribution of surface CO_2 exchange. This dissertation applies a geostatistical approach to inversions, which relies on an estimated spatiotemporal covariance structure to infer fluxes directly at fine scales in both space and time. In addition, process-based datasets can be incorporated into the inversion in a manner analogous to multi-linear regression, improving flux

estimates and providing inference regarding significant flux drivers. In the first dissertation component, environmental datasets are incorporated into a global inversion, with results showing that Leaf Area Index and the Fraction of Photosynthetically Active Radiation explain a significant portion of biospheric flux variability, while Gross Domestic Product and Population Density are associated with the fossil fuel emission signal. However, at the continental scale, flux estimates were found to be constrained primarily by the atmospheric measurements, with the grid-scale environmental datasets having minimal impact. The second component investigates the optimal use of continuous, continental CO_2 measurements influenced by the biospheric diurnal cycle, heterogeneous land-cover, and point-source fossil fuel emissions. In a series of synthetic data inversions over North America during the growing season, explicitly estimating the diurnal variability of fluxes was found to be critical for inferring unbiased fluxes at the aggregated monthly, ecoregion-scale. In the third component, a North American regional inversion is implemented using real data available from the continuous monitoring network in 2004. The biospheric portion of estimated total CO_2 flux is compared to a collection of bottom-up, process-based model output. Results show some convergence in the spatial patterns, seasonal cycle and net annual CO_2 flux between the inversion and bottom-up models, although inversion results at robust scales also help to provide insight into the forward model spread.

CHAPTER 1

Introduction

1.1 Why study the carbon cycle?

Atmospheric concentrations of carbon dioxide (CO_2) are currently growing at the rate of about 2 ppm per year due to fossil fuel combustion and land-use change (Canadell et al., 2007). In fact, the exponential increase since 1750 in atmospheric CO_2, a greenhouse gas, has contributed the strongest positive radiative forcing (among other anthropogenic factors) to global climate change since the start of the industrial era (Forster et al., 2007). This accumulation of CO_2 in the atmosphere is unprecedented in the last 650,000 years on Earth as shown in the ice core record (Siegenthaler et al., 2005), but it has also been partly counteracted by natural sinks in the terrestrial biosphere and oceans. An understanding of the processes driving these sinks and their future magnitude is relatively unknown, although this knowledge is essential for designing policies to stabilize atmospheric CO_2 concentrations.

Carbon exchange between the earth surface and the atmosphere can be divided roughly into three main components: land biosphere fluxes, ocean fluxes, and emissions from fossil fuel combustion, as shown in Figure 1.1 reproduced from the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (Denman et al., 2007). First, the terrestrial biosphere continuously exchanges CO_2 with the atmosphere through photosynthesis and respiration. Photosynthesis by plants fixes CO_2 from the atmosphere, and uses water and sunlight to convert this CO_2 into stored biomass. Respiration by plants (autotrophic respiration) or other organisms that directly consume biomass (heterotrophic respiration) subsequently return fixed CO_2 back to the atmosphere. These two biospheric processes on land are together termed Net

Ecosystem Exchange (NEE). While the magnitude of the photosynthesis and respiration fluxes are large compared to their residual on an annual basis (i.e. NEE), ecosystems can be net sources or sinks of carbon on multi-year timescales due to disturbances such as storms, fires or insect outbreaks that change the relative proportion of living vs. dead biomass, anthropogenic perturbations such as the CO_2 and nitrogen fertilization effects, or changing temperature and precipitation patterns from climate change itself (Denman et al., 2007). Fires, of both natural and anthropogenic origin, also release carbon back to the atmosphere directly through combustion. A related component of the carbon cycle is land-use change, particularly the clearing of forests for agricultural land-uses, which leads to reduced uptake capacity as well as large sources from decomposing biomass.

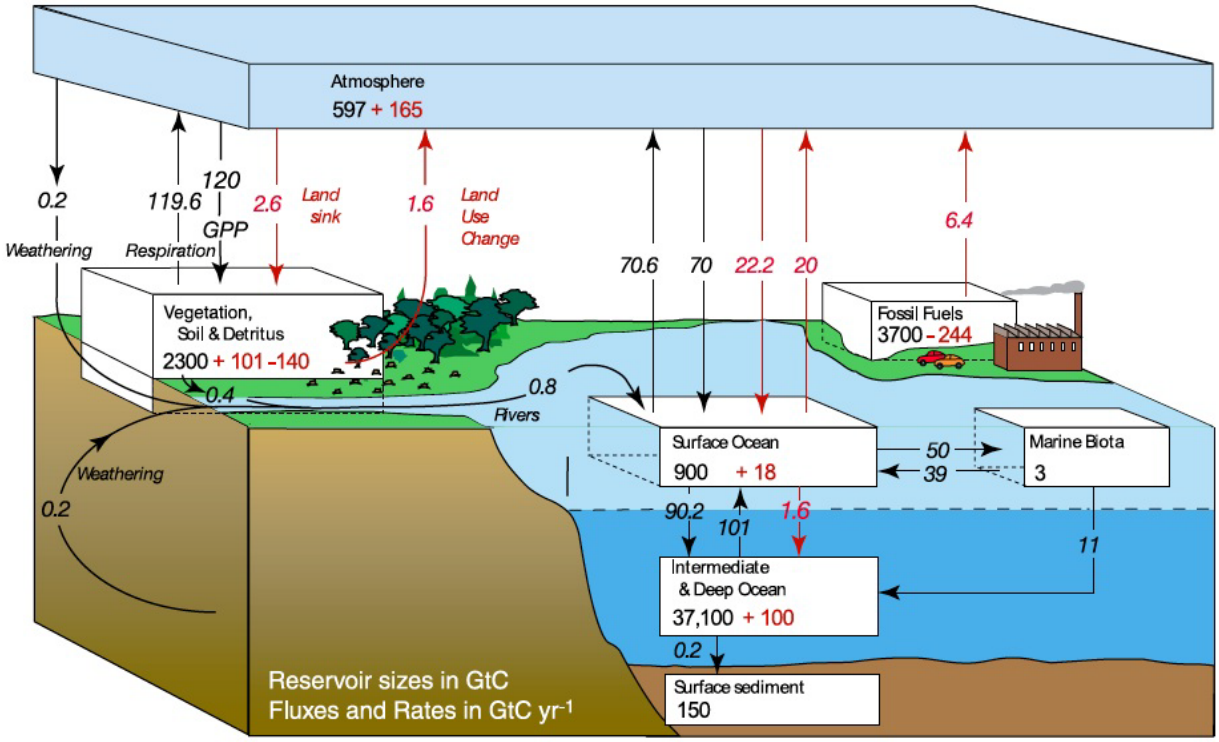

Figure 1.1: The magnitude of major carbon pools and fluxes on Earth in the industrial era. The black numbers represent "natural" pools and fluxes, whereas the red numbers represent changes induced by anthropogenic activities since 1750 (Source: IPCC AR4, Denman et al., 2007).

The oceans form the second main component of the carbon cycle, which exchange CO_2 with the atmosphere through air-sea gas transfer. Some of the dissolved CO_2 in the surface waters is taken up and used for photosynthesis by phytoplankton, which can die and sink, or in turn be consumed by other marine organisms. Dissolved and particulate carbon can then be exported to deeper waters through ocean circulation patterns, i.e. the solubility pump, and the sinking of dead organisms, i.e. the biological pump. The oceans can also be net sources of CO_2 to the atmosphere, with warmer waters leading to stronger out-gassing (Sarmiento & Gruber, 2006).

Finally, the last component, and the one of highest policy relevance, is CO_2 emissions due to fossil fuel combustion from oil, coal and natural gas. While fossil fuel emissions represent a small magnitude of flux in the context of the much larger individual gross fluxes between the atmosphere and the oceans and land biosphere, their net annual source to the atmosphere is large in comparison to the residual biospheric fluxes (Figure 1.1). Also, given the large magnitudes of the individual flux components of the terrestrial biosphere and oceans, small changes in their functioning, e.g. due to feedback effects from a changing climate, can lead to large changes in atmospheric CO_2 (Friedlingstein and Prentice, 2010). Therefore, in order to design policies to mitigate climate change, it is essential to have a better understanding of the carbon cycle in its entirety.

For the purposes of supporting climate change mitigation policies and improving mechanistic models for future prediction of the carbon cycle, it is important to be able to quantify current carbon exchange at the earth surface at relatively large regional scales, e.g. at the scale of political entities like states or provinces. Unfortunately, CO_2 fluxes cannot currently be directly measured at any scale larger than the approximately $1km^2$ footprint of an eddy-covariance flux tower (described in more detail in Section 1.2). Therefore, CO_2 fluxes at larger spatial scales can only be provided by models. Such models, which can be divided into top-down and bottom-up approaches, currently have a large spread in their estimates (see e.g., Figures 1.2 and 1.3), and reconciling these

model estimates provides the primary means of accurately quantifying regional-scale CO_2 exchange given the lack of any direct validation data.

1.2 Bottom-up understanding of CO_2 fluxes

The term "bottom-up" refers to a process-based understanding of CO_2 surface exchange, as gained from experimental studies at plot-level or laboratory scales. Bottom-up models of CO_2 flux may also include extrapolated relationships based on inventory data from forest growth measurements (e.g. Potter et al., 2008) and agricultural crop yield statistics (West et al., 2010), or even fossil fuel emissions from fuel sales and air pollution measurements (e.g. Gurney et al., 2009). In general, the spatial and temporal distribution of fossil fuel emissions is considered to be better-known than the biospheric component of the land carbon cycle (Marland et al., 2009), although in order to separately identify the biospheric and anthropogenic signals in the total CO_2 flux to the atmosphere, it is critical to have a better understanding of the terrestrial biosphere. Process-based bottom-up models also exist for the oceans, but consistent with the main focus of this dissertation, the following discussion will mostly focus on the terrestrial biosphere.

The only direct estimates of land-atmosphere CO_2 exchange, at scales larger than a laboratory, are from eddy-covariance flux towers. These flux estimates are derived from the covariance between continuously measured atmospheric CO_2 concentrations and vertical wind velocity (Baldocchi et al., 1988, 2003) within the footprint of the tower, and they provide rich datasets for testing hypotheses concerning the drivers of net biospheric carbon exchange at the land surface (Baldocchi, 2008). A growing network of these sites around the world (i.e. FLUXNET, Baldocchi et al., 2001; Friend et al., 2007) has also helped to increase understanding of the drivers of NEE in different climatic zones and ecosystems. However, due to heterogeneity in land cover and climatic gradients that vary on large spatial scales, it has proven difficult to extrapolate eddy-covariance CO_2 flux estimates to larger regions, although attempts have been made (e.g. Xiao et al., 2008; Jung et al., 2009). Also, these towers tend to be sited in

areas relatively remote from human influence, thereby inducing inherent biases in scaling their estimates to areas with managed ecosystems and large sources from fossil fuel emissions.

Bottom-up biospheric models estimate CO_2 fluxes associated with NEE by scaling up process-based understanding gained from experimental studies at smaller scales. Scaling up process-based understanding of NEE is as difficult as scaling up direct eddy-covariance estimates of CO_2 flux, due to heterogeneity in land cover and emergent processes at larger scales (Loureau et al., 2003; West et al., 2009). In fact, estimates of NEE can differ strongly between biospheric models due to varying assumptions about vegetation structure and input datasets (Cramer et al., 1999; Kicklighter et al., 1999), processes included in the model formulation, and overall level of complexity. For example, the Simple Biosphere Model or SiB (Sellers et al., 1996a,b)) mechanistically characterizes the energy and water cycles in addition to the carbon cycle, whereas more simple models such as the Vegetation Photosynthesis & Respiration Model (VPRM) rely on empirical relationships derived from flux tower data and remote-sensing products to simulate NEE (Mahadevan et al., 2008).

Another important distinction between biospheric models is whether they use values from remote sensing datasets to calibrate their estimates (diagnostic models), or they calculate variables like Leaf Area Index internally (prognostic models) (Huntzinger et al., in prep). While prognostic models are not as reliant on actual measurement data at the regional scale as are diagnostic models, they instead provide predictive ability which can be used for future scenario analysis, critical for evaluating interactions between the carbon cycle and a changing climate. For example, a class of biospheric models termed Dynamic Vegetation Models (e.g. Cramer et al., 2001; Krinner et al., 2005) attempt to simulate changes in plant functional types as environmental conditions change. Unfortunately, while these models can reproduce past carbon budgets reasonably well within levels of current understanding, they still diverge strongly in terms of their predicted response to future climate change (Sitch et al., 2008).

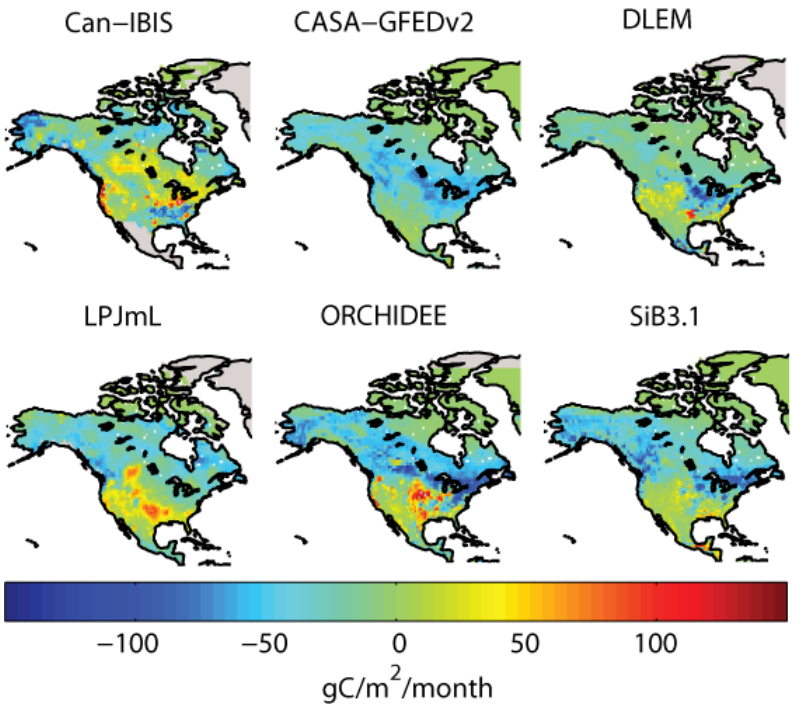

Figure 1.2: Example of the range of bottom-up estimates of CO_2 flux over North America during from the growing season from June to August, 2002 (Source: D.N. Huntzinger, personal communication).

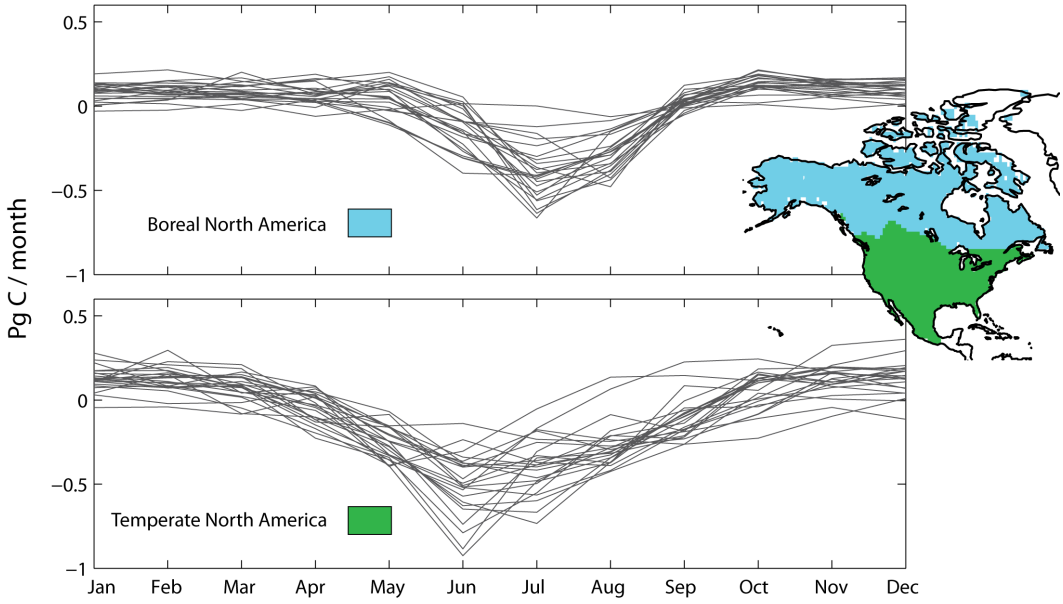

Figure 1.3: Example of the spread in the estimated long-term (2000-2005) seasonal cycle of NEE from 20 atmospheric inversions over North America. (Source: D.N. Huntzinger, personal communication).

1.3 Top-down approaches to studying the carbon cycle

In contrast to bottom-up approaches, "top-down" approaches for studying the carbon cycle attempt to scale down the signal associated with the net effect of all CO_2 exchange processes as seen in the atmosphere. Atmospheric CO_2 measurement time series from remote sites, like Mauna Loa, Hawaii, are helpful for understanding the seasonal and inter-annual variability of global-scale net CO_2 accumulation in the atmosphere (e.g. Keeling et al., 1995; Jones & Cox, 2005; Buermann et al., 2007). However, multiple measurement sites must be used in conjunction with an atmospheric transport model to deduce carbon sources and sinks at sub-global, i.e. continental or finer regional scales. Atmospheric inverse models, which are described in more depth throughout this dissertation, are used to accomplish this task.

Inverse modeling techniques, or top-down approaches, trace back variations in measured atmospheric CO_2 concentrations to the most likely configuration of surface sources and sinks with the aid of an atmospheric transport model (Enting, 2002). CO_2 is a long-lived, relatively inert gas, so that inverse problems for CO_2 sources and sinks do not need to consider atmospheric chemistry. However, the signal of surface flux variability becomes diluted over time due to atmospheric mixing, such that most information about the spatial distribution of sources and sinks is contained within the measurements only up to about six months after emission (Bruhwiler et al., 2005). Atmospheric mixing also tends to make CO_2 inverse problems ill-conditioned, in that there is a range of surface flux estimates that can reproduce the available measurement data (Enting, 2002). Finally, given the limited network of in-situ surface measurements typically used in CO_2 inversions (Tans & Conway, 2005), the problem may also be under-determined if the number of estimated fluxes is greater than the number of available observations.

In an ill-conditioned and/ or under-determined 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 (Fan et al., 1999). Bayesian inversion methods provide a form of regularization in which explicit prior flux estimates are included in the objective function (Enting, 2002). Therefore, solving a Bayesian inversion involves a compromise between estimating fluxes that can reproduce the atmospheric measurements when transported forward into concentration space, while also staying close to the prior flux estimates, particularly in under-constrained regions. This approach allows the inverse model to assimilate all relevant knowledge about carbon cycle science at the desired flux resolution.

The Bayesian objective function for an atmospheric inversion is typically formulated as:

$$L_{\mathbf{s}} = \frac{1}{2}(\mathbf{z} - \mathbf{Hs})^T \mathbf{R}^{-1}(\mathbf{z} - \mathbf{Hs}) + \frac{1}{2}\left(\mathbf{s} - \mathbf{s_p}\right)^T \mathbf{Q}^{-1}\left(\mathbf{s} - \mathbf{s_p}\right) \quad (1)$$

where $\mathbf{z}$ is a vector representing the atmospheric CO_2 measurements, $\mathbf{H}$ is the sensitivity matrix relating measured concentrations to surface fluxes ($\mathbf{s}$) in space and time, as derived from an atmospheric transport model, and $\boldsymbol{s_p}$ is a vector of prior flux estimates which are updated using the atmospheric data constraint. Two covariance matrices ($\mathbf{R}$ and $\mathbf{Q}$) provide information to the inversion as to how well it should be able to reproduce the concentration data given measurement, transport model and other sorts of errors in the model setup ($\mathbf{R}$), and how much to trust the prior flux estimates ($\mathbf{Q}$). Overall, by minimizing the objective function, an inversion balances these two competing goals of minimizing the mismatch between measured ($\mathbf{z}$) and "modeled" concentrations (i.e. $\mathbf{Hs}$, or the inferred fluxes transported forward to the measurement space), while also remaining faithful to the prior flux estimates ($\mathbf{s_p}$).

Atmospheric CO_2 inversions have until now principally been limited by the lack of data coverage in both space and time. The first major time series of atmospheric CO_2 concentration measurements was due to the efforts of Charles David Keeling, who first began collecting data at Mauna Loa, Hawaii in the 1950's, and began to document the

rise in atmospheric CO_2 over time from fossil fuel combustion (Harris, 2010). Since then, the NOAA-ESRL Global Monitoring Network (Tans and Conway, 2005) has grown to a global network of more than 100 sites in 2010 (shown in http://www.esrl.noaa.gov/gmd/dv/site/site_table2.html) that can be used to help constrain continental-scale CO_2 flux estimates in atmospheric inversion studies. However, most sites in this network provide only weekly flask samples, and the network still lacks spatial coverage in the tropics and southern hemisphere. Also, the flask sampling sites are primarily in remote or high-altitude locations, limiting their ability to inform sub-continental scale flux estimates.

In order to infer fluxes using an atmospheric inversion at sub-continental scales, the scale of political entities, it becomes necessary to use observation sites in the inversion located directly within high CO_2 flux variability areas (Law et al., 2003). For this purpose, there is a growing network of sites collecting continuous measurements in sub-continental, low-altitude locations. Some of these sites are also part of the NOAA-ESRL network, although again they are located primarily in North America and Europe (Bakwin et al., 1998; http://ce-atmosphere.lsce.ipsl.fr/database/index_database.html).

While the use of continuous, continental measurement data in inversions is promising for constraining flux estimates at finer spatial scales, the use of these data is particularly problematic during the growing season when incoming shortwave radiation drives both atmospheric dynamics and CO_2 flux processes. If not properly accounted for within the inversion framework, these co-varying processes can lead to biased flux estimates due to the diurnal rectifier effect (Denning et al., 1996). A related point is that in order to appropriately take advantage of continuous, continental data in an inversion, improvements in transport models are necessary to properly account for complex, small-scale dynamics in the near-field of the measurement locations (e.g. Lin et al., 2003). Meso-scale transport models remain an active area of research (e.g. Geels et al., 2007; Sarrat et al., 2007), not specifically discussed in this dissertation, that must advance in parallel to improvements in inverse model setup in order to appropriately take advantage of continuous, continental data-streams. Finally, the sheer volume of

the continuous measurement data and computational demands associated with meso-scale transport models that resolve transport dynamics on finer spatial scales has necessitated regional inversions that estimate fluxes only for a specific spatial domain, e.g. a continent.

While the availability of atmospheric CO_2 measurements and the reliability of the atmospheric transport model are the two components that critically underpin the quality of an atmospheric inversion, other setup choices can also affect the ability of the inversion to infer high quality flux estimates. These choices include the spatial and temporal resolution of inferred fluxes, the choice of prior flux estimates, the parameterization of covariance matrices, and the specification of boundary conditions for regional inversions with a limited domain. These setup choices are discussed in more detail throughout this dissertation.

1.4 Why the geostatistical approach to atmospheric inversions?

Ultimately, bottom-up and top-down methods for CO_2 flux estimation each have their advantages, and in many ways these are complementary approaches. Bottom-up models may be more accurate at finer scales than inversions, and they also provide the ability to help understand the processes driving the carbon cycle through sensitivity analyses of driving data and model formulation (e.g. Jain et al., 2005; Bondeau et al., 2007; Baker et al., 2008). In addition, these models provide the only means of predicting future carbon fluxes and expected carbon cycle-climate change feedback effects (Friedlingstein et al., 2006; Sitch et al., 2008). In contrast, the top-down approaches represent the carbon cycle from the point of view of the atmosphere, and at large scales, they may point to missing processes or errors in model formulation in mechanistic models (e.g. Gurney et al., 2003; Peylin et al., 2005). Particularly for monitoring fossil fuel emissions, top-down approaches have the potential to help validate fossil fuel emission inventories or identify missing sources in these datasets (e.g. Levin & Karstens, 2007; Djuricin et al., 2010; Pisso et al., 2010). With future global CO_2 data-streams from satellites, atmospheric inversions may also provide a means of

policing and enforcing CO_2 emission reduction commitments around the globe (NRC, 2010).

While bottom-up and top-down models have mostly complementary strengths and purposes, an ability to reconcile their CO_2 flux estimates at intermediate regional scales (e.g. individual states, provinces, or countries) would lead to increased confidence in results from both sets of models (Canadell et al., 2010). Currently, there exists a large spread in bottom-up model estimates of CO_2 flux, even for past years (e.g. Figure 1.2). However, there also exists a spread in inverse model estimates at both continental and sub-continental scales due to differing model setups and input data (e.g. Figure 1.3). With improvements in inverse modeling capabilities, the subject of this dissertation, there is a hope that the atmospheric constraint can be used to help validate forward models, or at least distinguish between inconsistent sets of bottom-up flux estimates. Better estimates of NEE from forward models can then help to more appropriately isolate the anthropogenic signal in the total CO_2 flux seen from the atmosphere to help enforce future policy commitments.

This dissertation makes a contribution towards helping to reconcile bottom-up and top-down understanding of regional-scale CO_2 fluxes by focusing primarily on the atmospheric constraint, and helping to disentangle the relative contribution of CO_2 measurements and bottom-up models used as explicit priors in previous synthesis Bayesian inversion studies (e.g. Rayner et al., 1999, 2008; Gurney et al., 2003; Rödenbeck et al., 2003; Baker et al., 2006). It does this by applying a geostatistical inverse modeling approach, which avoids the use of prior flux estimates (Michalak et al., 2004) by taking advantage of spatiotemporal autocorrelation in the estimated fluxes, thereby reducing the degrees of freedom in the solution. Also, this autocorrelation allows the inversion to solve for fluxes at relatively finer scales, thereby reducing scale-dependent errors in the solution (e.g. Kaminski et al., 2001; Tolk et al., 2008). Furthermore, the covariance parameters in the inversion are estimated with the Restricted Maximum Likelihood (Kitanidis, 1995; Michalak et al., 2004) method directly

using the atmospheric measurements, additionally reducing the reliance on bottom-up models or process-based assumptions for parameterizing the inversion.

While a geostatistical inversion can be run with a simple mean flux defined as the "prior", allowing for a completely independent comparison with bottom-up model output, process-based datasets can also be incorporated into the inversion to help downscale and extrapolate the atmospheric data constraint to under-sampled areas. These datasets can furthermore provide inferences regarding significant CO_2 flux drivers and their relationship to total CO_2 flux from the point of view of the atmosphere.

Overall, by relying more strongly on the atmospheric data constraint relative to other synthesis Bayesian inversion studies, the geostatistical approach helps to shed light on the impact of various setup choices that have a strong impact on flux estimates from an inversion, thereby helping to reduce the spread in results from other inversion approaches. Also, when not incorporating any process-based datasets, geostatistical inversions provide a completely independent comparison to bottom-up model output, making the results particularly useful for bottom-up/ top-down inter-comparison studies of CO_2 flux.

1.5 Objectives of this dissertation

This dissertation furthers the development of the geostatistical atmospheric inversion framework in a manner designed to help illuminate the impact of inversion setup choices on final flux estimates, and also close the gap between bottom-up and top-down understanding of CO_2 flux at regional scales. It does this by relying on the atmospheric data constraint as strongly as possible to parameterize the inversion, and carefully assessing potential sources of bias in the inversion setup that may have gone unnoticed in previous synthesis Bayesian inversion studies due to their reliance on explicit prior flux estimates from bottom-up models and other process-based assumptions.

By improving the quality of regional-scale flux estimates from an inversion and increasing their relative independence from process-based assumptions, the inversion

framework presented here is useful for providing insight into the spread of results from mechanistic forward models. As mentioned previously, the bottom-up forward models help to increase our understanding of the global carbon cycle through scenario analysis, while also providing predictive ability for carbon – climate feedback effects. However, their development has been hampered by the lack of direct validation data for regional-scale CO_2 exchange. By providing a framework for generating a robust set of estimates from an inversion primarily reliant on the atmospheric data constraint, this dissertation addresses this need for validation, potentially helping to overcome this limitation.

Finally, the geostatistical inversion provides the ability to help test and validate various process-based model assumptions and formulations by incorporating environmental datasets directly into the inversion with a regression-like approach. While these datasets can also help to improve flux estimates, particularly in areas under-constrained by the atmospheric measurements, they may be most useful for assessing the quality of bottom-up datasets or for testing process-based hypotheses concerning CO_2 flux drivers.

In the first major component of the dissertation (presented in Chapter 4, previously published as Gourdji et al. (2008)), the geostatistical inversion approach is applied at the global scale using atmospheric measurements from the NOAA-ESRL Cooperative Air Sampling Network to demonstrate the performance of the method with real data, relative to existing synthesis Bayesian inversions estimating global CO_2 fluxes for the same time period. Auxiliary environmental variables from remote-sensing and socioeconomic datasets are also incorporated into the inversion in a statistically rigorous manner to provide flux inference regarding significant global flux drivers, and assess the impact of including process-based information into the model on flux estimates at various spatial scales.

The second component of this dissertation (presented in Chapter 5, previously published as Gourdji et al. (2010)), drills down to the continental scale, using a regional inversion specifically for the North American continent. This component develops a regional inversion framework for optimally taking advantage of highly variable

continuous, continental measurement data. This task is accomplished through the use of pseudo-data inversions for June 2004, where the "true" solution is known and can be compared to inferred flux estimates within their recovered uncertainties. Two topics are specifically addressed: first, the use of the RML approach with atmospheric measurements is investigated for inferring appropriate covariance parameters for the inversion. Second, the impact of the flux temporal resolution is assessed, particularly whether or not to resolve diurnal variability in the estimated fluxes.

The third and final component of this dissertation (presented in Chapter 6), applies the setup developed in the second component with real measurements to infer CO_2 sources and sinks for 2004 over the North American continent. Additional challenges associated with using real measurement data, which could not be addressed in a synthetic data study, are addressed here, specifically data filtering to avoid systematic transport model errors, and the appropriate set of regional boundary conditions to use and their impact on final flux estimates. Without including any process-based datasets in the inversion, results are compared within recovered uncertainties to a collection of biospheric model flux estimates submitted to the North American Carbon Program regional interim synthesis (Huntzinger et al., in prep) for inter-comparison studies. Despite the limited measurement network for this year (9 continuous tower locations unevenly located across the continent and available flask and aircraft measurements), this inter-comparison helps to shed light on both the large spread of biospheric model results available for this year, as well as the strengths and weaknesses in the presented inversion setup and input data.

Chapter 2 provides a literature review of previous studies relevant to the presented work, while Chapter 3 describes the methods used. Chapters 4 to 6 contain descriptions of the individual components of the dissertation, while Chapter 7 concludes the dissertation and suggests promising avenues for future work.

In summary, this dissertation provides a contribution to the atmospheric inversion field by demonstrating the impact of various assumptions and setup choices on the final flux estimates, important for informing the robust use of atmospheric CO_2

concentration measurements in future operational carbon monitoring systems. Also, by increasing the reliance on the atmospheric data constraint to infer covariance parameters and final flux estimates, the presented inversions provide a more independent comparison to bottom-up forward models, thereby helping to close the gap between bottom-up and top-down understanding of CO_2 flux at regional scales.

CHAPTER 2

Literature Review

This chapter summarizes the state of the science in regards to the three main components of this dissertation: global atmospheric inversions using CO_2 measurements from the NOAA-ESRL Global Monitoring Network, regional atmospheric inversions using continuous data, and bottom-up/ top-down inter-comparisons of the North American carbon cycle. The literature review associated with the first component primarily covers methodological studies with some discussion of scientific results, while the discussion of the second component is entirely methodological. The discussion of the literature associated with the third component focuses on scientific conclusions that have emerged thus far primarily through bottom-up studies of CO_2 flux over the North American continent.

2.1 Global CO_2 inversion studies using flask measurements from the NOAA-ESRL Cooperative Air Sampling Network

Weekly flask samples of CO_2 concentration are collected at more than 100 sites world-wide in the NOAA-ESRL Cooperative Air Sampling Network (Figure 2.1; Tans & Conway, 2005), where these measurement locations have specifically been chosen to ensure sampling of well-mixed "baseline" air that is not influenced by local flux variability (Ramonet & Monfray, 1996). Many of these sites sample air from the Marine Boundary Layer (MBL) on islands or coastlines, while other sites are at high-elevation or in relatively remote locations. The flask sampling network has expanded from about 20 sites in 1980 to more than 100 locations world-wide today.

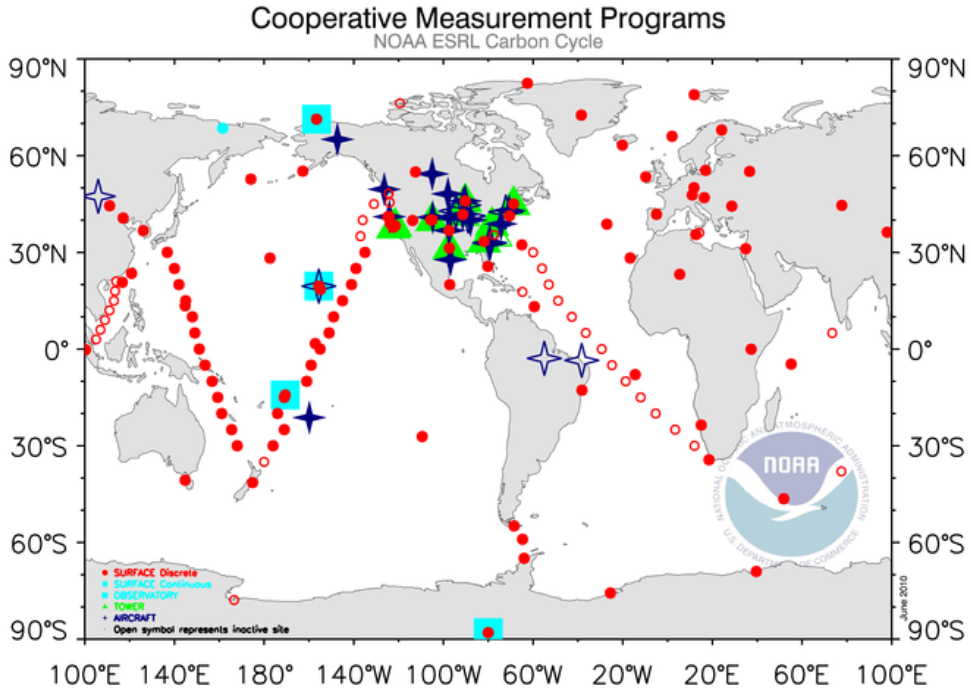

Figure 2.1: Map of measurement locations in the NOAA-ESRL Cooperative Air Sampling Network. Flask sampling sites are shown in filled red circles. (Source: http://www.esrl.noaa.gov/gmd/ccgg/).

Early atmospheric CO_2 inverse models used mass balance constraints and an observed gradient in atmospheric CO_2 concentrations between the northern and southern hemispheres to conclude that the "missing sink" in the global carbon budget was located in the Northern Hemisphere terrestrial biosphere (Tans et al., 1990; Ciais et al., 1995). Since then, a series of studies have used monthly atmospheric CO_2 observations derived from the global flask sampling network to estimate continental-scale fluxes in synthesis Bayesian inversions (e.g. Gurney et al., 2003, 2004; Law et al., 2003a; Baker et al., 2006; Rödenbeck et al., 2003), although global latitudinal bands still remain better-constrained than individual continents (Gurney et al., 2008).

Synthesis Bayesian CO_2 inversions typically specify biospheric model output, fossil fuel and fire emission inventories, and extrapolated ocean ship-track data as their prior flux estimates, and then estimate corrections at the scale of large regions to these

explicit priors (e.g. Bousquet et al., 2000; Gurney et al., 2002; Rödenbeck et al., 2003). Other studies using flask samples from the NOAA-ESRL Monitoring Network have taken the approach of using CO_2 measurements and the inversion framework to update parameters within a mechanistic forward model as part of a data assimilation system (Rayner et al., 2005; Scholze et al., 2007). While such studies maintain the predictive ability of the forward models and their partitioning of net CO_2 flux into various components (i.e. GPP, R_{a+h}, fires, etc.), most of the subsequent discussion will focus on synthesis Bayesian inversions that estimate CO_2 flux directly.

The spread in flux estimates from global synthesis Bayesian inversion studies associated with different data choices, transport models, covariance matrix structure and parameterization, prior flux estimates and flux estimation resolution have led to numerous methodological studies aiming to pinpoint optimal inversion setups, or more generally the impact of each of these input and setup choices on inversion results. Specifically, the TransCom3 series of inversion studies compared the performance of 17 different atmospheric transport models in constraining fluxes for 22 large regions around the globe, and concluded that transport model errors contribute a significant amount of uncertainty to the final flux estimates (Gurney et al., 2003, 2004; Law et al., 2003a; Baker et al., 2006). These coordinated studies also investigated the sensitivity of final flux estimates to the choice of priors, observing networks and other aspects of inversion setup. Relevant results will be referenced in the discussion below.

Limitations in the quality of atmospheric transport models are considered to be one of the largest sources of uncertainty in flux estimates from global inversions, with most of the differences observed in *a posteriori* fluxes from different transport models due to varying rates of inter-hemispheric and vertical mixing (Gurney et al., 2003). In fact, systematic errors in the vertical mixing from global transport models may have led to an overestimation of the sources in the tropics, and the sinks in the northern hemisphere (Stephens et al., 2007) from the TransCom3 and other global inversion studies (e.g. Gurney et al., 2002; Baker et al., 2006; Rödenbeck et al., 2003). Other studies have found that numerical problems in well-known transport models, apart from

errors in process-based formulation, can lead to very different scientific conclusions from global inversions, although these problems can be reduced somewhat by using finer spatial resolutions in the driving meteorology (Prather et al., 2008).

Given that atmospheric CO_2 inversions tend to be under-constrained, most studies using global flask samples have tried to take advantage of as much available data as possible. However, limitations of the transport model to appropriately simulate the dynamics around specific sites, coupled with sites coming in and out of the network over time, make it difficult to use all possible data in the inversion without inadvertently biasing flux estimates. For example, Patra et al. (2006) showed the impact on a global inversion of using ocean-only sites vs. the additional inclusion of land sites, where the transport is presumably more difficult to model. Not only did the spread in results using different transport models become wider after including the land-based sites, the flux over Europe changed from a neutral flux to a sink of 0.5 PgC/yr. Similarly, the introduction of a land-based low-altitude observation location in Canada to the TransCom3 global inversions reduced the source over Boreal North America from 0.25 PgC/yr to a neutral flux (Yuen et al., 2005). It is unclear if the large sensitivities found in these experiments to the inclusion of land-based sites represented a real signal in the measurements, or the impact of transport model errors associated with the simulation of atmospheric dynamics around these sites. Rödenbeck et al. (2003) similarly found that the choice of sites to include in the inversion can lead to geographical biases in the regions where particular measurement locations are included or excluded.

Another research area for data selection in global inversions concerns whether or not to use a smoothed data product (i.e. GlobalView, Masarie and Tans (1995), available from http://www.esrl.noaa.gov/gmd/ccgg/globalview/index.html) which can help to gap-fill small breaks in the record or filter out local flux variability that is difficult to model on the coarser grid of the transport model. While some inversion studies have directly used data from the GlobalView product in the inversion (e.g. Baker et al., 2006; Deng et al., 2007), other studies have used un-smoothed data in order to avoid losing signal in the actual measurements (e.g. Rodenbeck et al., 2003; Mueller et al., 2008;

Gourdji et al., 2008), although this may lead to representation errors due to the limited ability of the transport models to correctly model near-field variability in the data (Ciais et al., 2010). In order to take advantage of un-smoothed measurements in a global inversion, but eliminate measurements subject to systematic transport or representation errors, Maki et al. (2010) recently proposed a new filtering algorithm based on iterative runs of the inversion to remove data-points inconsistent with the flux signal seen by the network as a whole. Other studies have focused on the impact of sites coming into and out of the network over time, with Rödenbeck et al. (2003) concluding that it is best to only include measurement stations with complete data records over the course of the inversion when analyzing temporal trends in flux.

In addition to transport model errors associated with a given set of data choices, other errors can be introduced into the inversion by the choice of spatial and temporal resolution of the estimated fluxes or corrections to the prior (Engelen et al., 2002). These errors, termed aggregation errors, occur when the atmospheric measurements are sensitive to variability in the flux distribution at finer scales than the allowable corrections. Kaminski et al. (2001) specifically explored the impact of spatial aggregation errors on global inversions by quantifying errors in a synthetic data setup associated with resolving the globe using 18 vs. 54 regions. The results using only 18 regions had substantially higher errors than that using 54, with these errors being as large in magnitude as the fluxes themselves. Overall, Kaminski et al. (2001) recommended that global inversions should resolve fluxes at as fine a spatial scale as computationally feasible (with the limit being the resolution of the transport model). Peylin et al. (2002) focused on temporal aggregation errors, and showed that giving the inversion freedom to adjust monthly fluxes is preferable to estimating an annual correction to a seasonal cycle fixed from the prior. To help reduce the impact of aggregation errors, more recent global inversion studies have estimated fluxes at finer spatial and temporal scales by relying on an assumed correlation between grid-scale fluxes to help regularize the solution (e.g. Rödenbeck et al., 2003; Michalak et al., 2004; Mueller et al., 2008, Gourdji et al., 2008; Chevallier et al., 2010).

The choice of covariance parameters to use in an inversion has also been explored in the global inversion literature. Most studies have relied on analyses of the variability in CO_2 concentration measurements or modeled fluxes (e.g. Engelen et al., 2002; Gurney et al., 2004; Chevallier et al., 2006) in order to choose model-data mismatch and flux covariance parameters *a priori*. However, mis-specified covariance parameters can lead to biased *a posteriori* flux estimates and/ or under-estimated uncertainties resulting from the inverse model (Engelen et al., 2002; Chevallier et al., 2006; Gerbig et al., 2006). As an alternative to selecting covariance parameters based on analyses of the data or priors, a Maximum Likelihood approach was proposed by Michalak et al. (2005) as an objective way to estimate covariance parameters using the atmospheric data constraint within a Synthesis Bayesian inversion framework.

Despite the concern with the quality of inversion setup and input data, some robust scientific conclusions have emerged from global inversion studies using *in situ* flask measurements. These include the impact of El Niño/ La Niña and volcanic eruptions like Mount Pinatubo in 1991 on the inter-annual variability of global CO_2 exchange that can be clearly seen from the global flask sampling network. For example, net land sinks were seen to increase following the Mt. Pinatubo eruption in 1991 in both the tropics and northern hemisphere extra-tropics (Gurney et al., 2008), although mechanistic explanations for this sink still vary (Farquhar & Roderick, 2003; Angert et al., 2004). Most of the inter-annual variability in global CO_2 fluxes was also found to be driven by the land biosphere in the tropics and southern latitudes due to El Niño/ La Niña effects (Rödenbeck et al., 2003; Baker et al., 2006).

Studies aiming to pinpoint more accurately the spatial location of net sources and sinks around the globe from inversion studies have been less conclusive. One interesting result from the TransCom3 studies (Gurney et al., 2002), which was consistent with other global inversion studies as well (e.g. Peylin et al., 2002; Rödenbeck et al., 2003), was that the ocean sink in the Southern Ocean appeared to be weaker than oceanographic measurements would suggest (Takahashi et al., 2002). In fact, later work by Takahashi et al. (2009) using extrapolated ocean ship-track measurements of pCO_2

revised earlier estimates to be more in agreement with inversion results. However, the distribution of the Northern Hemisphere terrestrial land sink still varies somewhat among different inversions, with some studies locating this sink mainly in North America (e.g. Fan et al., 1998) and others in Eurasia (Bousquet et al., 1999). Butler et al. (2010) recently showed significant redistributions of the Northern Hemisphere terrestrial sink across Asia, Europe and North America using different measurement networks in a global inversion, while another recent bottom-up/ top-down synthesis study comparing four global inversions with land-based carbon accounting data (Ciais et al., 2010) argues that the largest proportion of this sink is in Russia.

Most of the global inversion studies discussed above used the Synthesis Bayesian approach to estimate fluxes, where spatial patterns and fine-scale temporal variability were fixed from the explicit prior flux estimates. This is known to cause significant aggregation errors in the flux solution (e.g. Kaminski et al., 2001; Peylin et al., 2002). Therefore, the need for approaches that can incorporate process-based information into the inversion in a manner consistent with the atmospheric data constraint, while simultaneously reducing the impact of aggregation errors by estimating fluxes at fine spatial and temporal scales, is the primary motivation for the first major component of this dissertation (Chapter 4).

2.2 Regional CO_2 inversion studies using continuous, continental measurement data

The well-mixed air sampled by the weekly flasks from the NOAA-ESRL global monitoring network limits their ability to constrain sub-continental scale terrestrial sources and sinks (Gloor et al., 2000). However, an expanding tower network in North America and Europe calibrated to international standards (Bakwin et al., 1998; http://ce-atmosphere.lsce.ipsl.fr/database/index_database.html), many of these also NOAA-ESRL sites, continuously samples atmospheric CO_2 concentrations in continental, low-altitude areas with high local flux variability. The CO_2 mixing ratio measurements from these towers at certain times of the day contain information about regional-scale

fluxes (~1000 km x 1000 km) (Gloor et al., 2001; Bakwin et al., 2004) if used in conjunction with an atmospheric transport model that can adequately simulate atmospheric dynamics in the near-field of the measurement sites (Geels et al., 2007). Also, in contrast to the weekly measurements used in global inversions, inversions using continuous data may be capable of inferring fluxes at even finer spatial scales if the transport models can appropriately take advantage of the information contained in the fine temporal variability of measured concentrations (Law et al., 2002).

The use of continuous CO_2 concentration measurements from continental, low-altitude measurement locations in an inversion can be potentially problematic, given the strong diurnal cycle of flux in heavily-vegetated areas, the covariance between fluxes and the height of the atmospheric boundary layer and the inability to correctly model this height in atmospheric transport models (i.e. the "diurnal rectifier effect" (Denning et al., 1996)). Other sorts of errors associated with modeled transport in the near vicinity of the tower, particularly point-source fossil fuel emissions that are difficult to resolve at the scale of the driving winds, may also complicate the use of this data in inversions. In fact, the use of continuous, continental measurements in an inversion puts strong demands on the quality of the transport model in order to infer realistic surface flux estimates. Towards this end, a new class of meso-scale meteorological models has helped to better describe fine-scale continental atmospheric dynamics within a limited domain (e.g. Sarrat et al., 2007). Lagrangian transport models, coupled with the meso-scale meteorological models, can then interpolate to a point-based measurement (e.g. Uliasz and Pielke, 1991; Lin et al., 2003), thereby avoiding larger representation errors associated with the use of coarse-scale global Eulerian models (e.g. TM3, Heimann and Körner, 2003).

It should be noted that the use of coupled Lagrangian and meso-scale meteorological models can help to improve transport accuracy in the near-field of continental measurements, but at the cost of computational expense. Resolving fluxes at finer spatial and temporal scales to avoid aggregation errors also necessitates a means to reduce computational costs. Therefore, regional inversions that estimate

fluxes for only a limited domain have emerged as a means to take advantage of high-quality transport models while estimating fluxes at fine scales. Unfortunately, regional inversions require the specification of CO_2 concentration boundary conditions at the edge of the domain, which can introduce additional uncertainties into the inversion calculations (Gerbig et al., 2006). In fact, errors in boundary conditions used in regional real-data inversions have been shown to introduce a significant bias into flux estimates, particularly at the annual timescale (Peylin et al., 2005; Göckede et al., 2010b). More recent studies have attempted to develop a framework for nesting a Lagrangian transport model for a specific regional domain within a larger global inversion using a coarser Eulerian model (e.g. Rödenbeck et al., 2009; Trusilova et al., 2010). This setup allows one to estimate fluxes and model transport at finer scales within the region of interest, while simultaneously eliminating the need for boundary conditions.

Apart from concerns with the transport model and boundary conditions, an appropriate inversion setup is also important for optimally taking advantage of continuous, continental CO_2 measurements in an inversion (e.g. Carouge et al., 2010b). Synthetic data (a.k.a. "pseudo-data") experiments are particularly useful for diagnosing inversion quality because they include a set of specified baseline fluxes with which results can be compared, making it easier to diagnose potential biases in inferred fluxes under a number of different scenarios (Law et al., 2002). Also, synthetic data studies can help to isolate the impact of inversion setup choices, as opposed to other sources of bias in inversions, e.g. due to errors in the boundary conditions, transport model, or prior flux estimates. For example, the effect of atmospheric transport model errors can be controlled by using the same transport model to create the synthetic measurements as is used to estimate fluxes in the inversion, and covariance parameters can also be derived from the "true" underlying flux distribution (Gourdji et al., 2010, or Chapter 5 of this dissertation).

In a series of pseudo-data inversions, Law et al. (2002, 2003b, 2004) investigated the optimal inversion setup for taking advantage of continuous measurements collected on or near the Australian continent. The first of these studies (Law et al., 2002) used 4-

hourly synthetic measurements from marine boundary layer locations. Spatial aggregation errors were shown to be particularly important when using continuous data in this study, consistent with the results from Schuh et al. (2009). In a follow-up study, Law et al. (2003b) performed similar synthetic data inversions, but varied the averaging intervals of the concentration data from 4-hourly to monthly. The study concluded that data averaging at the synoptic scale (2-5 days) helped to reduce the impact of transport model errors, but that in general, bias in the recovered fluxes grew with coarser-scale concentration averaging intervals. While these first two studies used monthly average fluxes to generate the synthetic concentration data, Law et al. (2004) tackled the more realistic assumption of a diurnally and synoptically-varying underlying flux field. Not surprisingly, this study found that it was necessary to explicitly resolve the diurnal cycle in the estimated fluxes in order to avoid biases due to the diurnal rectifier effect (Denning et al., 1996).

Carouge et al. (2010a, b), investigated a regional inversion using a 10-tower continuous network over the European continent in a pseudo-data environment. This study found a more reliable constraint on fluxes in the western portion of the continent where measurement locations are most dense. She also found that results from the inversion were most robust when post-aggregated to 1000 km x 1000 km and 10-day averaging intervals, while results at finer scales were highly sensitive to inversion setup.

Pseudo-data experiments can also help to assess the constraint on fluxes that is achievable using an idealized setup. For example, Gerbig et al. (2006) showed in a pseudo-data scenario that inversions that take advantage of continuous, continental CO_2 measurements, high-resolution transport modeling, and diagnostic biospheric models based on remote-sensing inputs for prior flux estimates, can reduce uncertainties on recovered fluxes at the continental scale by 2-3 orders of magnitude compared to coarse-scale global inversions. However, it is still unclear if this uncertainty reduction is achievable in a real data environment.

For example, in a real-data environment, limitations of the transport model necessitate the use of measurements from only certain times of the day. Geels et al.

(2007), in an inter-comparison study of regional atmospheric transport models over Europe, concluded that only afternoon values at low-altitude sites can be represented sufficiently well by the current suite of transport models, and that only these values should be used for constraining sources and sinks in regional inversions. In fact, the use of only afternoon data has become the norm in most regional inversion studies (e.g. Peters et al., 2007; Schuh et al., 2010; Butler et al., 2010). While the well-mixed air during this time ensures that these measurements are the most regionally representative (Haszpra, 1999; Bakwin et al., 2004), the use of afternoon measurements alone may limit the ability of the inversion to infer fluxes at relatively fine spatial scales (Mueller et al., in prep.).

In the first real-data regional inversion over Europe, Peylin et al. (2005) estimated daily average fluxes for November 1998 using daily-averaged concentration data from 6 continuous measurement locations. By estimating daily-averaged fluxes, this study avoided dealing with their underlying diurnal variability. However, given that November in the Northern hemisphere has a reduced diurnal cycle compared to months in the height of the growing season, this study assumed that biases due to temporal aggregation errors could be adequately accounted for in the model-data mismatch covariance matrix. While noting the strong influence of *a priori* spatial covariance assumptions on flux estimates, this study also concluded that the influence of far-field fluxes grew throughout the inversion, and that an accurate set of boundary conditions was critical for correctly estimating regional fluxes.

Matross et al. (2006) used concentration measurements from a tall tower in Argyle, Maine in a real-data regional inversion to estimate scaling factors on photosynthesis and respiration fluxes for different vegetation types in the New England and Québec region. This study concluded that the atmospheric data from a single tower were not able to distinguish well between different vegetation types within the domain and that therefore aircraft data and/or towers with overlapping footprints were necessary to better constrain regional carbon budgets in future inversions.

A number of inversion studies now estimate sub-continental scale fluxes over the entire North American continent using real continuous measurement data. The CarbonTracker data assimilation system (Peters et al., 2007), designed and maintained by the National Oceanic & Atmospheric Administration (NOAA), estimates fluxes using afternoon-average continuous data from measurement towers within North America, while simultaneously running a global inversion to provide boundary conditions for the inflow of air into the North American continent. The Peters et al. (2007) study estimated weekly net fluxes from 2000 to 2005 for 25 eco-regions in North America using a synthesis Bayesian approach embedded within an ensemble Kalman filter data assimilation framework. Similarly, Butler et al. (2010) used a Synthesis Bayesian global inversion including data from continuous measurement sites in North America to estimate monthly fluxes for 10 sub-continental regions within the continent. Both CarbonTracker and the Butler et al. (2010) study fixed grid-scale spatial patterns from the bottom-up prior fluxes, and then adjusted these patterns at coarser scales. Schuh et al. (2010) used a regional inversion setup to estimate fluxes over North America using weekly, grid-scale (1°x1°) bias corrections to individual photosynthesis and respiration flux components from the prior.

Given the need to understand the strengths and limitations of using continuous CO_2 concentration data in inversions and their interactions with inversion setup, most studies taking advantage of this data to the current time have been methodological in nature, as discussed above. Scientific conclusions have mostly been tentative given the strong sensitivity to inversion setup seen in these previous studies. However, firm methodological conclusions include that when using continental, continuous measurement data in inversions, there is a need to resolve fluxes at fine scales to avoid representation and aggregation errors, and that accurate transport models are important for simulating dynamics in the near-field of the measurement locations. The second component of this dissertation (Chapter 5) makes a contribution to this literature by extending the work from the first component (Chapter 4), using the geostatistical inverse modeling framework, to investigate scale-dependent errors

associated with using continental, continuous measurement data collected over North America. The third component of this dissertation (Chapter 6) uses the lessons learned from the first two components regarding inversion setup to resolve fluxes at the finest scale that is computationally feasible, using the best available transport model, to recover flux estimates over North America.

2.3 North American carbon cycle science

The principal purposes of running regional CO_2 inverse models are 1) to evaluate, to the extent possible, biospheric models and fossil fuel inventories and provide a means to distinguish between the assumptions and scaling parameters contained within these process-based models, and 2) to provide estimates of carbon balance at various spatial scales and partitioned into relevant processes for informing carbon management policies. In order to properly interpret flux estimates from an atmospheric inversion and help achieve both of these objectives, it is necessary to understand the state of the science in regards to the carbon cycle in the domain and scale of interest. The third component of this dissertation estimates fluxes over North America using measurement data for 2004, and therefore the following review will focus specifically on North American carbon cycle science, as inferred principally from inventory and bottom-up modeling studies.

Overall, North America is a large source of carbon to the atmosphere, with 1.75 GtC in 2004 from fossil fuel emissions alone (Gurney et al., 2009; Oda and Maksyutov, 2010). However, many studies show that the North American land surface is also a net biospheric sink, helping to counteract roughly 30% of fossil fuel emissions on an annual average basis, as shown in an inventory-based approach from the State of the Carbon Cycle Report focusing on North America (SOCCR; CCSP, 2007). Ecosystems on the continent were estimated to absorb about 0.67 GtC/yr in 2003 in this study, of which 0.16 GtC/yr was exported outside the region by rivers and international trade. Since the great majority of this exported carbon is returned to the atmosphere within a year, the overall sink for this year was estimated at 0.51 GtC/yr (CCSP, 2007).

The North American biospheric sink, which can vary from year to year, has been variously attributed to fire suppression and woody encroachment in arid rangelands, reduced logging, increased agricultural productivity, no-till agriculture, the re-growth of formerly agricultural land in the eastern United States, excess nitrogen deposition from air pollution, CO_2 fertilization, and climatic changes, i.e. wetter conditions and longer growing seasons (Pacala et al., 2001; Nemani et al., 2002; CCSP, 2007). While roughly half of the overall sink in North America can be attributed to re-growth of formerly agricultural land, the uncertainties on the magnitude of the sink due to the other causal factors are still about 100% (CCSP, 2007). Mexico remains a carbon source to the atmosphere from ongoing deforestation (Masera et al., 1997; Velasquez et al., 2010).

Although the biospheric sink in the North American continent during the 1980's and 1990's was remarkably stable (Pacala et al., 2001), this sink is expected to slow down as re-growing forests reach maturity (CCSP, 2007). However, much uncertainty still remains regarding the future of the North American biospheric carbon cycle. Current research focuses on the impact of inter-annual climatic variability (e.g. Desai et al., 2010) and long-term mean changes on the net sink. For example, the 2002 drought on the continent in which 45% of the United States was classified as "Extreme" or "Exceptional" in the U.S. Drought Monitor (http://drought.unl.edu/dm) reduced carbon uptake to approximately 0.3 GtC/yr, only half of the long-term mean (Peters et al., 2007). Longer growing seasons due to climate change enhance net uptake in the spring and fall, but higher temperatures can also increase rates of evapotranspiration leading to water stress, thereby limiting productivity in ecosystems not currently considered to be water-limited (e.g. Xiao & Moody, 2004; Welp et al., 2007). The interacting effects of CO_2 fertilization and increasing nitrogen deposition in temperate forests, which are typically considered to be nitrogen-limited, are also still unclear (Finzi et al., 2006; Pregitzer et al., 2008).

In general, the agricultural portion of the North American carbon cycle is even less well-understood at large scales than that of forests and other ecosystems in their natural, pre-industrial state, although recent modeling and inventory-based studies

have helped to address this gap (e.g. Lokupitiya et al., 2009; Corbin et al., 2010; West et al., 2010). The Mid-Continent Intensive Campaign (MCI, http://www.nacarbon.org/nacp/mci.html) of the North American Carbon Program also has the aim to investigate carbon cycling over the Midwestern agricultural regions in the United States. While preliminary work failed to find convergence between crop inventory, bottom-up models and inversion approaches, more recent work is beginning to show an increased convergence between different estimates (A. Schuh, personal communication). Overall, research in the MCI region should provide insight into the net flux associated with different crop types, planting schedules and potentially increased planting for biofuel production (Searchinger et al., 2008).

Changes in human settlement patterns and land management practices will all impact the North American biospheric carbon cycle in the coming century. For example, a projected net increase of about 6 million hectares in pine plantations in the southern U.S. will lead to increased carbon uptake, although increasing urban sprawl and land conversion from forest and agriculture to human settlements will counteract this trend (Pataki et al., 2006; Alig and Butler, 2004). Also, proper incentives for carbon sequestration could increase the use of no-till agriculture and enhanced forestry practices, although more research is needed to determine the efficacy of various approaches (Baker et al., 2007).

The most promising means of restoring the North American land surface to a neutral carbon balance for climate change mitigation is to reduce fossil fuel emissions across the continent, most likely from a combination of factors including a switch to alternative energies, increased energy efficiency of buildings, cars and industries, and higher-density residential development (Pacala and Socolow, 2004). In terms of scientific understanding, fossil fuel emissions in North America are considered to be better-known than the biospheric portion of the total CO_2 flux (Marland et al., 2009); however, active research also exists to refine the space-time distribution of bottom-up fossil fuel inventories over the continent, principally for use in atmospheric CO_2 inversion studies. Initial fossil fuel inventories were mostly at the annual time-scale and

were based on assumptions of how country-level fuel consumption varies with population (e.g. Andres et al., 1996). Later work showed that there is a strong seasonal cycle in fossil fuel emissions, particularly in continental climates with heating and cooling seasons (Gurney et al., 2005), which can have a strong impact on inversions that pre-subtract the influence of fossil fuel emissions from the atmospheric measurement data *a priori*. Traffic patterns and residential behavior can also create a diurnal cycle of emissions. The Vulcan dataset has been recently developed to address these concerns by estimating fossil fuel emissions over the continental United States at an hourly, 10 km resolution by relying on data from fuel sales, power plant locations and air quality reporting to the Environmental Protection Agency over the last few decades (Gurney et al., 2009). Monthly fossil fuel emission inventories have also been created for the globe (e.g. Oda and Maksyutov, 2010), which can help to fill in seasonal variability in emission estimates over North America in areas excluded by the Vulcan inventory.

In addition to using inventories and bottom-up models, the North American carbon balance has also been estimated using the atmospheric data constraint. For example, the inferred continental carbon budget from CarbonTracker was an average biospheric sink of 0.65 PgC/yr from 2000 to 2005, in close agreement with the above-mentioned inventory-based SOCCR estimate of 0.67 GtC/yr in 2003 (before export from the continent due to rivers and international trade; CCSP, 2007). Similarly, Crevoisier et al. (2010) used a novel approach for carbon budgeting based on the inflow and outflow of air from the free troposphere using measurements of the vertical gradient of atmospheric CO_2 across the continent, and found a biospheric sink of 0.5 PgC/yr from 2004 to 2006. Crevoisier et al. (2010) also determined from this analysis that the greatest uptake occurs in the agricultural Midwest and the Southeast. While CarbonTracker also places net uptake in the agricultural areas, the Southeast is relatively under-sampled in the in-situ CO_2 monitoring network, and therefore the CarbonTracker results are not consistent with Crevoisier et al. (2010) in this region.

A few studies to date have performed top-down/ bottom-up inter-comparisons on the North American continent to elucidate process-based assumptions in the

bottom-up models. These studies have mainly focused on the boreal areas (e.g. Dargaville et al., 2002; McGuire et al., 2010) where there is a strong interest in predicting the response of boreal ecosystems to projected future climate change. Relying on top-down estimates from the TransCom3 studies (e.g. Gurney et al., 2004), Dargaville et al. (2002) showed that the atmospheric constraint implied a stronger peak uptake shifted a month later as compared to four different biospheric models estimating flux in Boreal North America. McGuire et al. (2010) drew similar conclusions for a study in the Arctic basin. The inversion results also showed stronger inter-annual variability in these studies than the biospheric models, implying missing or incorrect process formulations, perhaps associated with soil freeze-thaw dynamics.

The third component of this dissertation (Chapter 6) makes a contribution to the bottom-up/ top-down inter-comparison literature over North America by extending the analysis to all biomes and land cover types across the continent, comparing inversion results to a collection of 16 different biospheric models with a wide spread in their estimates of NEE. This inter-comparison relies on results from the regional geostatistical inversion over North America implemented using an optimal setup and lessons learned from the first two components of this dissertation, thereby providing improved inversion quality at finer spatial scales than previously used global inversion results. Also, because the geostatistical inversion eliminates the requirement for explicit prior flux estimates from biospheric models, flux estimates are relatively independent of the forward models, providing a unique opportunity to shed light on their spread at various spatiotemporal scales.

CHAPTER 3

Methods

This chapter contains a survey of the methods implemented in this dissertation. Please note that generalized versions of the Matlab code that was used to implement each of the presented techniques can be downloaded from http://puorg.engin.umich.edu.

3.1 Geostatistical inverse modeling

Given atmospheric mixing that dilutes the influence of surface sources and sinks of CO_2, atmospheric inversions must introduce additional information into the model to help constrain flux estimates, other than the direct CO_2 concentration measurements and transport model sensitivities relating measurements to surface fluxes (Enting, 2002). Inversions are therefore typically formulated as a Bayesian problem where the atmospheric measurements are used to update prior estimates of the surface flux distribution, where these prior estimates are primarily derived from bottom-up model output (e.g. Gurney et al., 2003). This dissertation uses a modified Bayesian approach to CO_2 flux estimation termed geostatistical inverse modeling that does not require prior estimates of the magnitude and spatial distribution of surface CO_2 fluxes (Michalak et al., 2004). Instead, an assumed spatial and/ or temporal correlation is introduced between flux estimates to help constrain the solution. Process-based datasets related to CO_2 flux may additionally be incorporated into the inversion to provide inference and further constrain flux estimates, if desired.

In the geostatistical approach, the flux distribution is modeled as the sum of a deterministic but unknown component, $\mathbf{X\beta}$, referred to as the linear trend of the

surface flux distribution, and a zero-mean stochastic component with a spatial and/or temporal autocorrelation described by the covariance matrix $\mathbf{Q}$. The linear trend ($\mathbf{X\beta}$) implemented here replaces the explicit prior flux estimates used in a Synthesis Bayesian inversion, with this trend defining the portion of the flux signal that can be explained by a set of covariates included in the matrix $\mathbf{X}$. The $\boldsymbol{\beta}$ values, or drift coefficients, corresponding to these covariates are then estimated using the atmospheric data. This spatiotemporal trend can be as simple as a constant mean flux, but can also include spatially and temporally-varying auxiliary variables with a process-based relationship to CO_2 flux. The overall best estimates obtained through this approach minimize deviations from the lienar trend, as well as residuals between actual atmospheric CO_2 measurements and concentrations derived from the estimated fluxes. In the discussion that follows, m represents the number of estimated fluxes, n is the number of atmospheric concentration measurements, and p is the number of covariates included within the trend.

The objective function $L_{\mathbf{s},\boldsymbol{\beta}}$ for a geostatistical inversion is defined as:

$$L_{\mathbf{s},\boldsymbol{\beta}} = \frac{1}{2}(\mathbf{z}-\mathbf{Hs})^T\mathbf{R}^{-1}(\mathbf{z}-\mathbf{Hs}) + \frac{1}{2}(\mathbf{s}-\mathbf{X\beta})^T\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{X\beta}) \qquad (3.1)$$

where the vector $\mathbf{z}$ ($n\times1$) represents the atmospheric CO_2 measurements (ppm), and $\mathbf{s}$ ($m\times1$) is the vector of fluxes (μmol/(m^2s)). $\mathbf{H}$ ($n\times m$) describes the sensitivity of CO_2 measurements to surface fluxes, as quantified from an atmospheric transport model, with units of ppm/(μmol/(m^2s)), and $\mathbf{Hs}$ therefore represents a vector of modeled CO_2 observations. $\mathbf{X}$ is a known ($m\times p$) matrix containing the flux covariates in the model of the trend, $\boldsymbol{\beta}$ are ($p\times1$) unknown drift coefficients, and $\mathbf{X\beta}$ is the resulting trend of CO_2 flux. The two covariance matrices in the objective function, $\mathbf{R}$ ($n\times n$) and $\mathbf{Q}$ ($m\times m$), balance the relative weight of the atmospheric data and the trend in estimating fluxes. $\mathbf{R}$ is the model-data mismatch covariance matrix, describing the expected magnitude of discrepancies between observed ($\mathbf{z}$) and modeled ($\mathbf{Hs}$) CO_2 concentrations (due to measurement, transport model, representation, and aggregation errors). $\mathbf{Q}$ ($m\times m$) is

the *a priori* flux covariance matrix, characterizing how flux deviations from the model of the trend (i.e. $\mathbf{s} - \mathbf{X\beta}$) are correlated in time and space.

The $\mathbf{X}$ matrix ($m \times p$) defines the covariates included within the trend. Each of the p covariates is defined at the time and location of each of the m estimated fluxes. $\boldsymbol{\beta}$ is a vector ($p \times 1$) of coefficients, estimated as part of the inversion, that correspond to the variables in $\mathbf{X}$ and represent the linear relationships between each variable and CO_2 flux, as seen through the atmospheric data. A flux covariate can be as simple as a column of ones and zeros where the associated $\boldsymbol{\beta}$ would represent a mean flux over the flux locations and time periods included as ones in $\mathbf{X}$. Alternatively, the flux covariates could be any spatially and/ or temporally-varying dataset with a process-based correlation to CO_2 flux, e.g. a vegetative index like Leaf Area Index or a climatological variable such as precipitation. The overall trend $\mathbf{X\beta}$ is conceptually similar to a multivariate linear regression where the components in $\mathbf{X}$ are predictor variables that explain some portion of the flux variability, and $\boldsymbol{\beta}$ are the coefficients on these variables. However, unlike multiple linear regression, the relationships are estimated in an inverse modeling framework (using concentration measurements to infer the coefficients on the covariates), and the approach assumes spatially and temporally-correlated deviations from the trend (as specified in $\mathbf{Q}$). In order to be consistent with terminology commonly used in statistics, the $\boldsymbol{\beta}$ values in this dissertation are referred to as drift coefficients.

By minimizing the objective function defined in equation (3.1) with respect to $\mathbf{s}$ and $\boldsymbol{\beta}$, the inversion simultaneously minimizes differences between the estimated fluxes ($\mathbf{s}$) and the model of the trend ($\mathbf{X\beta}$), and the residuals between actual atmospheric CO_2 measurements ($\mathbf{z}$) and concentrations derived from the estimated fluxes ($\mathbf{Hs}$). The $\mathbf{R}$ and $\mathbf{Q}$ covariance matrices control the balance between achieving these two objectives. For example, low variances in the model-data mismatch covariance matrix ($\mathbf{R}$) drive the inversion to reproduce the measurement data at the expense of keeping flux estimates close to the model of the trend. Also, in areas with more sensitivity to the available measurements, as described by $\mathbf{H}$ which is generated from the transport model, the inversion relies more heavily on reproducing observations, whereas in areas lacking

measurements, the inversion reverts more strongly to the model of the trend ($\mathbf{X\beta}$) and the spatiotemporal correlation of flux residuals ($\mathbf{Q}$).

Minimizing equation (3.1) with respect to fluxes, $\mathbf{s}$, and drift coefficients, $\boldsymbol{\beta}$, yields the following system of linear equations:

$$\begin{bmatrix} \boldsymbol{\Psi} & \mathbf{HX} \\ (\mathbf{HX})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Lambda}^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{HQ} \\ \mathbf{X}^T \end{bmatrix} \tag{3.2}$$

where

$$\boldsymbol{\Psi} = \mathbf{HQH^T} + \mathbf{R} \tag{3.3}$$

The weights $\boldsymbol{\Lambda}$ $(m \times n)$ and Lagrange multipliers $\mathbf{M}$ $(p \times m)$ are estimated by solving this system of linear equations, and are then used to define the estimated fluxes ($\hat{\mathbf{s}}$) and their posterior covariance ($\mathbf{V}_{\hat{\mathbf{s}}}$) as:

$$\hat{\mathbf{s}} = \boldsymbol{\Lambda}\mathbf{z} \tag{3.4}$$

$$\mathbf{V}_{\hat{\mathbf{s}}} = -\mathbf{XM} + \mathbf{Q} - \mathbf{QH}^T\boldsymbol{\Lambda}^T \tag{3.5}$$

where this full covariance matrix ($\mathbf{V}_{\hat{\mathbf{s}}}$) represents a composite of uncertainty associated with the estimation of unknown drift coefficients ($\boldsymbol{\beta}$) in the model of the trend, the spatiotemporal variability of fluxes as represented in $\mathbf{Q}$, and the overall constraint on fluxes as determined by the concentration footprints, the model of the trend and the prior covariance matrices.

Estimates of the drift coefficients, $\widehat{\boldsymbol{\beta}}$, and their uncertainty covariance ($\mathbf{V}_{\hat{\beta}}$) are calculated as:

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{Q}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Q}^{-1}\boldsymbol{\Lambda}\mathbf{z} \tag{3.6}$$

$$\mathbf{V}_{\hat{\beta}} = (\mathbf{X}^T\mathbf{H}^T(\mathbf{HQH}^T + \mathbf{R})^{-1}\mathbf{HX})^{-1} \tag{3.7}$$

where the diagonal elements of $\mathbf{V}_{\hat{\beta}}$ represent the uncertainties of the drift coefficients, and the off-diagonal terms in $\mathbf{V}_{\hat{\beta}}$ represent their error covariance.

The estimated fluxes ($\hat{\mathbf{s}}$) can also be expressed in a form more similar to that used in synthesis Bayesian inversions, as the sum of a deterministic component ($\mathbf{X}\widehat{\boldsymbol{\beta}}$), i.e. the estimated model of the trend of the flux distribution, and a stochastic component that is a function of the *a priori* correlation structure in $\mathbf{Q}$:

$$\hat{\mathbf{s}} = \mathbf{X}\widehat{\boldsymbol{\beta}} + \mathbf{Q}\mathbf{H}^T(\mathbf{H}\mathbf{Q}\mathbf{H}^T + \mathbf{R})^{-1}\left(\mathbf{z} - \mathbf{H}\mathbf{X}\widehat{\boldsymbol{\beta}}\right) \tag{3.8}$$

3.2 Covariance parameter optimization using Restricted Maximum Likelihood

As discussed previously, the *a priori* spatiotemporal covariance structure in $\mathbf{Q}$ helps to regularize a geostatistical inversion. To be precise, the covariance matrix $\mathbf{Q}$ represents the spatial (and/or temporal) correlation structure of flux residuals from the trend, and therefore this correlation depends on the degree to which the model of the trend $\mathbf{X}\boldsymbol{\beta}$ can represent the flux variability inferred using available observations. In a simple case where $\mathbf{X}$ is a column of ones representing a single mean flux in space and time, $\mathbf{Q}$ would describe the correlation of the fluxes themselves. At the opposite extreme where the model of the trend could reproduce all the inferred variability in fluxes, the flux residuals would be zero and $\mathbf{Q}$ would become a diagonal matrix. In practice, however, flux residuals are always non-zero and correlated due to imperfect prior information and process-based datasets available to represent flux in the trend, as well as weaknesses in the inversion framework itself (e.g. transport model error).

For all geostatistical inversions in this dissertation, the $\mathbf{Q}$ matrix is modeled using an exponentially decaying spatial correlation between flux residuals from the trend:

$$Q_{ij}\left(h_{ij}|\sigma^2, l\right) = \sigma^2 \exp\left(-\frac{h_{ij}}{l}\right) \tag{3.9}$$

where h_{ij} is the separation distance between two estimation locations. The practical correlation length is approximately $3l$, beyond which σ^2 represents the expected variance of independent flux residuals. The choice of an exponential decay was based on the work of Michalak et al. (2004), and on variogram analyses of bottom-up flux estimates.

In Chapter 5, temporal correlation is additionally considered between flux residuals and this becomes:

$$\boldsymbol{Q}(h_x, h_t|\sigma^2, l, \tau) = \sigma^2 exp\left(-\frac{h_x}{l}\right) \exp\left(-\frac{h_t}{\tau}\right) \quad (3.10)$$

where h_x and h_t are the separation distances between grid cells in space and the lag in time, respectively, τ is the temporal correlation range parameter, and σ^2 and *l* are as described above.

The model-data mismatch variances in the $\mathbf{R}$ matrix, include measurement, transport, aggregation and representation errors for each observation. Measurement errors are most typically calibration errors (e.g. Francey & Steele, 2003), while transport errors can be random or systematic, a more problematic and perhaps more realistic form of error, e.g. associated with misrepresentations of the height of the PBL. Aggregation and representation errors are closely related concepts that can both impact the ability of an inversion to infer realistic fluxes. Aggregation errors occur when fluxes are estimated at coarse spatial and temporal resolutions, when in fact the measurements are sensitive to finer-scale variability in fluxes that is inconsistent with the patterns assumed at coarser scales (Kaminski et al., 2001). Representation errors refer specifically to the mismatch between a point-based measurement of CO_2 and the coarser resolution of the transport model and driving meteorology for the gridcell containing the measurement (e.g. Tolk et al., 2008). Although some sensitivity tests were performed to investigate the impact of correlated model-data mismatch errors, results were inconclusive as to the benefit of accounting for this correlation. Therefore,

all of the presented inversions in this dissertation were performed with a diagonal $\mathbf{R}$ matrix.

While most inversion studies choose covariance parameters based on "expert knowledge" using assumptions about how model-data mismatch scales with temporal variability in the data and known transport model errors (e.g. Baker et al., 2006; Butler et al., 2010), and also by examining the variability of biospheric model output for the *a priori* flux covariance structure (e.g. Rödenbeck et al., 2003), this dissertation takes a more statistical and quantitative approach for estimating covariance parameters. The parameters of the $\mathbf{Q}$ and $\mathbf{R}$ matrices are optimized using the Restricted Maximum Likelihood (RML) method (Kitanidis, 1995; Michalak et al., 2004; Mueller et al., 2008), a quantitative approach that helps to reduce biases in the flux estimates associated with errors in the covariance matrices.

To estimate flux covariance parameters in $\mathbf{Q}$, RML can be implemented using a set of bottom-up flux estimates from biospheric models, fossil fuel and fire emission inventories and oceanic datasets. Alternatively, RML can be implemented in an inverse setup using the atmospheric measurements, allowing for simultaneous estimation of parameters in both $\mathbf{Q}$ and $\mathbf{R}$. In this dissertation, RML was implemented with bottom-up datasets for the global study (Chapter 4), and with atmospheric data for the North American inversions (Chapters 5 and 6), as discussed further in the individual chapters associated with this work. Equations for both setups are shown below.

To estimate optimized covariance parameters, the RML approach minimizes the negative log-likelihood of the available observations with respect to the covariance parameters ($\boldsymbol{\theta}$). When using a set of bottom-up flux estimates ($\mathbf{s}$) to estimate the flux covariance parameters in $\mathbf{Q}$ (i.e. $\boldsymbol{\theta} = \{\sigma_{\mathbf{Q}}^2, l, \tau\}$), the corresponding equation is (Kitanidis, 1995):

$$\mathrm{L}_{\boldsymbol{\theta}_{\mathbf{Q}}} = \frac{1}{2}\ln|\mathbf{Q}| + \frac{1}{2}\ln|\mathbf{X}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{X}| + \frac{1}{2}[\mathbf{s}^{\mathrm{T}}(\mathbf{Q}^{-1} - \mathbf{Q}^{-1}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Q}^{-1})\mathbf{s}] \qquad (3.11)$$

Alternatively, using the atmospheric measurements ($\mathbf{z}$) to estimate covariance

parameters rather than the bottom-up model output ($\mathbf{s}$), $\boldsymbol{\theta}$ contains both $\mathbf{Q}$ and $\mathbf{R}$ parameters (i.e. $\{\sigma_{\mathbf{Q}}^2, l, \tau\}$ and one or more $\sigma_{\boldsymbol{R}}^2$ values associated with various subsets of the measurement data), and this equation becomes:

$$\mathrm{L}_{\boldsymbol{\theta}} = \frac{1}{2}\ln|\boldsymbol{\Psi}| + \frac{1}{2}\ln|(\mathbf{HX})^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\mathbf{HX}| + \frac{1}{2}\left[\mathbf{z}^{\mathrm{T}}(\boldsymbol{\Psi}^{-1} - \boldsymbol{\Psi}^{-1}\mathbf{HX}((\mathbf{HX})^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\mathbf{HX})^{-1}(\mathbf{HX})^{\mathrm{T}}\boldsymbol{\Psi}^{-1})\mathbf{z}\right] \quad (3.12)$$

where $\boldsymbol{\Psi}$ replaces $\mathbf{Q}$, and $\mathbf{HX}$ replaces $\mathbf{X}$ in equation 3.11. The RML objective function is minimized using a Trust-Region Reflective Newton algorithm, available through Matlab's optimization toolbox (MathWorks, 2010).

By using the atmospheric measurements to estimate covariance parameters, one does not have to rely on modeled flux estimates to infer the covariance structure of the "true" distribution. However, atmospheric mixing and limitations of the transport model may also make it difficult to infer reasonable covariance parameters using this approach. Parameters in $\mathbf{R}$ can only be estimated by RML using the atmospheric measurements.

3.3 Statistical variable selection techniques

Two different types of statistical model selection techniques were implemented in this dissertation to choose flux covariates for the linear trend ($\mathbf{X}$). For the global inversion study (Chapter 4), the Variance-Ratio Test (Kitanidis, 1997) was implemented with modifications for an inversion setup. As discussed below, this is a hypothesis-based testing approach which allows for comparison between nested models only. To circumvent this limitation, the Bayesian Information Criterion (Schwarz, 1978), a criterion-based approach which can compare all possible models, was implemented using a geostatistical inversion setup for the 2004 real data inversion (Chapter 6).

The Variance-Ratio Test (Kitanidis, 1997) is a hypothesis-based variable selection method that was originally developed to justify the inclusion of a more complex trend in

geostatistical interpolation, i.e. kriging. A modified method, compatible with an inverse modeling setup, is presented and implemented here. These modifications were made in an analogous manner to the RML equations using the atmospheric data, i.e. substituting $\mathbf{HQH}^{\mathrm{T}} + \mathbf{R}$ for $\mathbf{Q}$, and $\mathbf{HX}$ for $\mathbf{X}$ in the original Kitanidis (1997) equations.

In a geostatistical inversion, improving the model of the trend's ability to represent CO_2 flux variability can increase the accuracy of the recovered flux distribution, and reduce the associated *a posteriori* uncertainty. However, adding auxiliary variables with only a spurious correlation to flux can bias the model, and yield unreasonable estimates in poorly constrained areas. The Variance-Ratio Test is designed to balance the risks of including too few versus too many variables, by quantifying the significance of the improvement in model fit resulting from the addition of one or more variables to the model of the trend.

In this approach, the Weighted Sum of Squares (WSS) of the orthonormal residuals is defined for an initial ($\mathbf{X}_0$, $m \times p$) and an augmented ($\mathbf{X}_1$, $m \times (p+q)$) model of the trend (where $\mathbf{X}_0$ is a subset of $\mathbf{X}_1$) as:

$$WSS = \mathbf{z}^T(\mathbf{\Psi}^{-1} - \mathbf{\Psi}^{-1}\mathbf{HX}(\mathbf{X}^T\mathbf{H}^T\mathbf{\Psi}^{-1}\mathbf{HX})^{-1}\mathbf{X}^T\mathbf{H}^T\mathbf{\Psi}^{-1})\mathbf{z} \qquad (3.13)$$

WSS is a measure of fit that assesses how well the two trends, $\mathbf{X}_0$ and $\mathbf{X}_1$, explain the variability in fluxes as seen through the atmospheric concentration measurements, $\mathbf{z}$, and as weighted by the appropriate covariance matrices ($\mathbf{R}$ and $\mathbf{Q}$). The WSS equation, as presented above, accounts for the spatial correlation of the residuals in order to create a test analagous to model selection for multivariate linear regression. The WSS equation (3.13) was also modified for an inversion setup from that presented in Kitanidis (1997) by substituting $\mathbf{HX}$ for $\mathbf{X}$ and $\mathbf{\Psi}$ for $\mathbf{Q}$.

A trend with more auxiliary variables will always be able to represent a greater amount of the inferred variability relative to a model with fewer variables. For example, WSS_1 is always less than or equal to WSS_0, given that $\mathbf{X}_1$ includes all the variables in $\mathbf{X}_0$, as well as one or more additional variables. However, adding too many variables runs

the risk of "overfitting" and thereby introducing spurious noise or unphysical relationships into the model. Therefore, in order to quantify the significance of the improvement in model fit, we calculate the normalized relative difference (v) between WSS_0 and WSS_1 as:

$$v = \frac{(WSS_o - WSS_1)/q}{WSS_1/(n-p-q)} \tag{3.14}$$

The significance level of this statistic is quantified using an F distribution with q and $n - p - q$ degrees of freedom (where n represents the number of available measurements, p the number of components in $\mathbf{X}_0$, and q the number of additional components in $\mathbf{X}_1$ relative to $\mathbf{X}_0$). Only variables with a significant improvement in model fit are included into the augmented model. The Variance-Ratio Test can then be repeated using the augmented model as $\mathbf{X}_0$ until there are no significant variables left to include in the trend.

In contrast to hypothesis-based tests like the Variance-Ratio Test, criterion-based tests compare all possible models, and therefore allow for comparison of non-nested models (Ward, 2008). For this dissertation, the Bayes Information Criterion (Schwarz, 1978) was chosen to select auxiliary variables for the 2004 real data inversion (Chapter 6). This method is generally preferred to other criterion-based approaches, e.g. the Akaike Information Criterion (Akaike, 1974), when the goal is inference rather than prediction, and the BIC tends to select fewer variables than the AIC.

The original BIC equations from Schwarz (1978) were modified for a geostatistical setup with correlated residuals (Mueller et al., 2010), as well as an inverse formulation using the atmospheric measurements. After these modifications, the criterion that must be minimized for the BIC approach can be expressed as:

$$BIC = \ln|\mathbf{\Psi}| + [\mathbf{z}^{\mathrm{T}}(\mathbf{\Psi}^{-1} - \mathbf{\Psi}^{-1}\mathbf{H}\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{\Psi}^{-1}\mathbf{H}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{H}^{\mathbf{T}}\mathbf{\Psi}^{-1})\mathbf{z}] + p\ln(n) \tag{3.15}$$

where *n* is the number of observations, and *p* represents the number of covariates within a given model ($\boldsymbol{X}$).

Given that comparing all possible models (2^p) quickly becomes computationally expensive and perhaps infeasible for large supersets of variables, the BIC was implemented with a Branch-and-Bound algorithm (Land and Doig, 1960) to help make the problem computationally tractable. This algorithm avoids unnecessary matrix multiplications by eliminating model "branches" as it runs that cannot possibly contain the "best" model (Yadav et al., in prep). The Branch-and-Bound algorithm relies on the fact that the Residual Sum of Squares within equation 3.15 is a monotonic function that can only decrease as variables are excluded from the model.

3.4 Geostatistical inversion algorithm

Running a complete geostatistical inversion first requires assembly of all input data in the proper format. Measurement data must be collected, processed and stacked into a linear vector ($\mathbf{z}$). Sensitivity matrices ($\mathbf{H}$) must be generated from the atmospheric transport model, and potential auxiliary variables for the model of the trend ($\mathbf{X}$) must be collected, (potentially) normalized, and trimmed to the specific scale and domain of interest. The structure of the covariance matrices must also be chosen *a priori*, e.g. whether to include both spatial and temporal flux covariance in $\mathbf{Q}$, or just spatial covariance, and whether to use an exponential decay or some other model. Also, one must decide how many separate model-data mismatch variance parameters to optimize based on an initial understanding of which measurement locations and/or times share similar data characteristics.

Once all input components have been assembled, running a geostatistical inversion requires three main steps (Figure 3.1), with the first two steps possibly requiring repetition until results stabilize. First, using an initial guess of the covariance parameters, the statistical variable selection methods (i.e. Variance-Ratio Test or BIC) must be run to choose covariates for the model of the trend ($\mathbf{X}$) that can optimally explain the signal in the atmospheric data. Second, the RML algorithm is run to

optimize covariance parameters associated with the selected covariates in $\mathbf{X}$. The variable selection step then needs to be repeated with the optimized covariance parameters to ensure that the selected variables do not change. If the variables do change, then the RML and variable selection steps must be repeated once again.

Finally, the inversion is run using the CO_2 observations ($\mathbf{z}$), sensitivity matrices ($\mathbf{H}$), selected covariates in $\mathbf{X}$, and covariance matrices ($\mathbf{R}$ and $\mathbf{Q}$) with optimized parameters, in order to estimate grid-scale fluxes ($\hat{\mathbf{s}}$) and their *a posteriori* covariance ($\mathbf{V}_{\hat{\mathbf{s}}}$), as well as the drift coefficients ($\hat{\boldsymbol{\beta}}$) associated with the trend and their associated covariance ($\mathbf{V}_{\hat{\boldsymbol{\beta}}}$). Flux estimates and uncertainties can then be post-aggregated to any coarser spatial and temporal scale of interest.

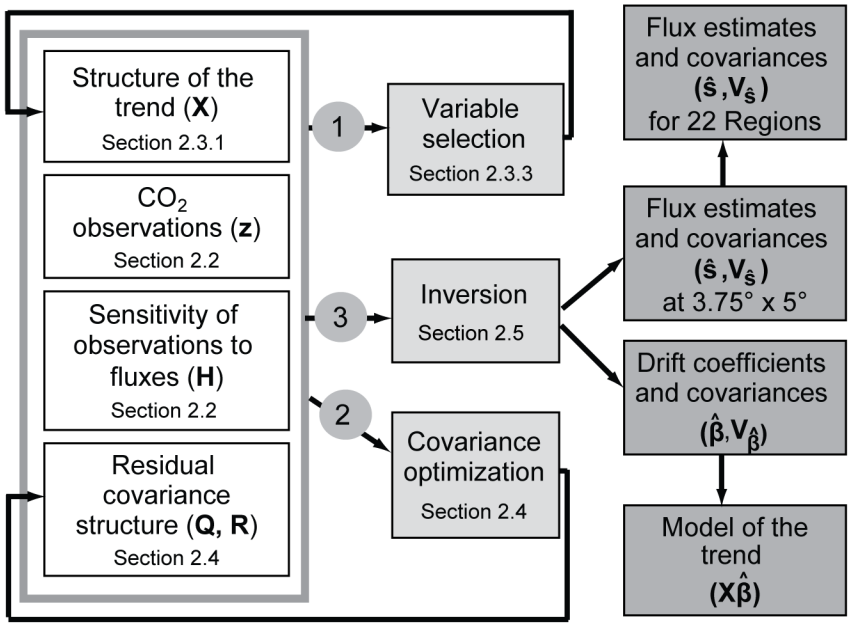

Figure 3.1: Schematic of geostatistical inversion components and algorithm. White boxes indicate inversion inputs, light gray boxes indicate inversion steps, and dark gray boxes represent inversion outputs. Grey circles indicate the sequence of steps in the algorithm.

CHAPTER 4

Using auxiliary environmental variables to help constrain grid-scale CO_2 flux estimates within a global geostatistical inversion

4.1 Introduction

Atmospheric inverse modeling, a technique for estimating CO_2 sources and sinks at the earth surface, uses atmospheric concentration measurements and an atmospheric transport model to estimate surface flux distributions. Given the currently limited network of atmospheric measurement locations, however, inverse problems aimed at CO_2 flux estimation are ill-posed and frequently under-determined. To circumvent these problems, most previous inverse modeling studies have used a synthesis Bayesian inversion approach, where *a priori* assumptions about both the magnitude and spatial patterns of fluxes are included in the inversion. This prior information is typically derived from biospheric model output, extrapolated ocean ship-track data, and fossil fuel inventories, and is then updated using atmospheric CO_2 observations (e.g., Kaminski et al., 1999; Rödenbeck et al., 2003; Gurney et al., 2004; Baker et al., 2006).

Geostatistical inverse modeling differs from these previous approaches by eliminating the need for explicit prior flux estimates, thereby allowing for more strongly atmospheric-data-driven estimates of global flux distributions (Michalak et al., 2004; Mueller et al., 2008). The geostatistical approach uses a modified Bayesian setup to estimate the flux distribution as the sum of a deterministic but unknown overall trend,

and a stochastic spatially and/or temporally autocorrelated flux residual. The trend in a geostatistical framework can be as simple as a global average land or ocean flux (Michalak et al., 2004), but can also include linear combinations of global auxiliary environmental datasets related to CO_2 flux. The covariates included in the model of the trend help to define the spatial and temporal structure of the flux distribution, but the exact relationship between these covariates and flux is estimated using the atmospheric CO_2 data themselves. In this way, available auxiliary data can help to constrain the inverse problem, but their impact on the flux distribution is not specified *a priori*. The stochastic component of a geostatistical estimate represents features of the flux distribution that are inferred from the CO_2 observations, but that cannot be explained using the covariates included in the model of the trend.

In a companion paper, Mueller et al. (2008) demonstrated the ability of the geostatistical approach to recover monthly grid-scale (3.75° x 5°) CO_2 fluxes using atmospheric concentration data from a subset of the NOAA-ESRL cooperative air sampling network (Tans and Conway, 2005). In that application, the trend was defined as monthly-varying land and ocean global average fluxes. Mueller et al. (2008) showed that the information content of available atmospheric measurements was sufficient to constrain fluxes at aggregated continental scales, particularly on land. Grid-scale estimates, however, had limited sub-continental spatial variability and high *a posteriori* uncertainties.

The primary objective of the current paper is to investigate the additional constraint provided by auxiliary environmental datasets on flux distributions estimated within a geostatistical inverse modeling framework. These datasets may include variables such as Leaf Area Index and Gross Domestic Product, which correlate well with the spatio-temporal pattern of biospheric and anthropogenic CO_2 exchange. Given their global coverage and correlation with CO_2 flux, these variables are also able to provide information about flux in regions under-constrained by the atmospheric measurements. A subset of auxiliary variables is selected based on their ability to explain flux variability evident from the available atmospheric data. The effect of

including auxiliary variables on *a posteriori* estimates and their associated uncertainties is investigated at two spatial (grid and continental) and two temporal (monthly and annual) scales. The goal is to provide a framework that allows the inversion to recover more realistic CO_2 flux variability with lower *a posteriori* uncertainties, relative to a setup relying exclusively on the limited atmospheric CO_2 measurement network.

The second objective is to investigate the relationship between the selected auxiliary datasets and flux, as identified using the atmospheric CO_2 observations, the uncertainty associated with this inferred model, and the impact of this uncertainty on the overall *a posteriori* uncertainty associated with the flux distribution. The relationships between each of the variables and the estimated fluxes are not pre-specified in the geostatistical inversion, but rather quantified using the atmospheric observations. If the environmental datasets are relatively objective quantities with global coverage, the inclusion of auxiliary variables in the inverse model can incorporate process-based information into the final flux estimates while minimizing assumptions about the relationship between the auxiliary datasets and CO_2 flux. Finally, given that the relationship between each variable and flux is estimated as part of the geostatistical inversion, these relationships can potentially be used to improve process-based understanding of flux drivers.

In this work, monthly CO_2 fluxes and their uncertainties are estimated for 1997 to 2001 at a 3.75° x 5° resolution for the globe. The model of the trend of the flux distribution incorporates auxiliary environmental variables, as discussed above, as well as monthly-varying terrestrial latitudinal flux gradients and global average ocean fluxes. The monthly latitudinal gradients, the direction and magnitude of which are estimated as part of the inversion, account for seasonal and latitudinal patterns in terrestrial fluxes that are not fully explained by the auxiliary variables. Results are compared to estimates presented in Mueller et al. (2008) using solely the atmospheric data constraint.

Note that the presented application estimates the total CO_2 flux, including the biospheric, anthropogenic and oceanic components. This is in contrast to previous

inversion studies which considered fossil fuel emissions well-known and estimated only the biospheric and oceanic portions of the flux distribution (e.g. *Rödenbeck et al.,* 2003; *Baker et al.,* 2006). By pre-subtracting a static dataset of fossil fuel emissions from the observational data, previous inversion studies aliased any spatial and temporal uncertainty in the fossil fuel flux distribution onto the biospheric fluxes or nearby ocean regions. Given that fossil fuel emissions, at least in the Northern Hemisphere, are known to vary seasonally, pre-subtracting assumed fossil fuel emissions can confound the interpretation of *a posteriori* fluxes (Gurney et al., 2005).

This chapter is organized as follows: Section 4.2 presents an overview of the inputs into the geostatistical inversion, with an emphasis on the auxiliary environmental data which are incorporated into the estimation. Section 4.3 presents the results of the analysis, including the selected auxiliary variables and their impact on flux estimates. Section 4.4 summarizes the main conclusions of the study.

4.2 Methods

The surface flux estimates presented in this paper are obtained using a geostatistical inverse modeling approach, minimizing the objective function in equation 3.1. A full description of this method is provided in Chapter 3 as well as in Michalak et al. (2004) and Mueller et al. (2008). This section presents a description of extensions to the method developed and implemented through the current work. A diagram of the overall algorithm can be found in Chapter 3, Figure 3.1.

4.2.1 Observational data (z) and transport model (H)

Monthly-averaged atmospheric CO_2 flask measurements ($\mathbf{z}$) from 44 unevenly-distributed global measurement locations within the NOAA-ESRL cooperative air sampling network (Tans and Conway, 2005) are used to constrain the global flux distribution, together with a transport matrix, $\mathbf{H}$, describing the sensitivity of measured concentrations to estimated fluxes. These components of the inversion are identical to those presented in the companion paper (Mueller et al., 2008). The observational

subset in $\mathbf{z}$ is similar to that used in Rödenbeck et al. (2003), and the number of measurements in any given month ranges from 35 to 42 between 1997 and 2001. The $\mathbf{H}$ matrix was derived from an adjoint implementation of the atmospheric transport model TM3 (Heimann and Körner, 2003), which has a spatial resolution of 3.75° latitude by 5° longitude with 19 vertical levels, and is driven by inter-annually varying winds from the NCEP Reanalysis (Kalnay et al., 1996).

4.2.2 Model of the trend $(\boldsymbol{X\beta})$

4.2.2.1 Structure of the model of the trend

The simple model of the trend implemented by Mueller et al. (2008) includes estimated average fluxes for each calendar month over land and ocean, and thereby captures both seasonal variability and differences in the expected flux magnitude over land and oceans. The model of the trend presented in the current study replaces these monthly average land fluxes with a subset of spatially and temporally-varying auxiliary environmental variables, selected using the procedure presented in Section 4.2.3.3. In addition, a monthly-varying terrestrial latitudinal gradient, expressed as $\sin(2 \times latitude)$, is included to represent the expected opposing sources and sinks in the Northern and Southern hemispheres. The strength and direction of this gradient is allowed to vary seasonally, in order to reflect the seasonality in the two hemispheres. A monthly-varying spatially-constant mean is assumed for ocean fluxes, identical to the setup used in Mueller et al. (2008).

Overall, the structure of the trend in this study is represented by an $(m \times (24 + k))$ matrix $\mathbf{X}$, where the first 24 columns contain the monthly terrestrial latitudinal flux gradients and ocean constants, and the subsequent k columns contain the auxiliary variables for each month and location:

$$\mathbf{X} = [\mathbf{A}_1 \quad \dots \quad \mathbf{A}_{12} \quad \mathbf{b}_1 \quad \cdots \quad \mathbf{b}_k] \tag{4.1}$$

where $\mathbf{b}_i$ includes values of the ith auxiliary variable for each of the m estimated

fluxes. The matrices $\mathbf{A}_j$ $(m \times 2)$ contain non-zero entries only for fluxes within a single calendar month j. For a given month, the relevant portion of this matrix, defined as the 3456×2 matrix $\mathbf{a}_j$, contains values of $\sin(2 \times latitude)$ for land gridcells in the first column, and ones for ocean gridcells in the second column:

$$\mathbf{a}_j = \begin{bmatrix} \sin(\mathbf{2} * \mathbf{latitude}) & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \quad (4.2)$$

4.2.2.2 Auxiliary environmental variables

The goal of incorporating auxiliary variables associated with carbon cycle processes into the model of the trend is to better represent the expected spatial and temporal variability of *a posteriori* grid-scale flux estimates, while only including variables that provide significant information as seen through the atmospheric monitoring network. A preliminary set of auxiliary variables with global coverage for the study period was selected based on the variables' known associations with biospheric or fossil fuel fluxes. Few oceanic variables with complete spatial and temporal coverage are available for 1997 to 2001. In addition, although the study initially considered variables such as Sea Surface Temperature, these were eliminated given preliminary results showing that the atmospheric data were not able to infer physically-reasonable relationships to flux in the oceans. As more oceanic datasets with gridded, global coverage become available, especially from the MODIS (Moderate Resolution Imaging Spectroradiometer) instrument on the Terra and Aqua satellites, future studies could make use of this information to better explain oceanic flux variability.

The auxiliary variables considered in this study are presented in the first column of Table 4.1, and described below. All variables were regridded from their native resolutions to the 3.75° x 5° resolution of this study using area-weighted averaging.

Downwelling Shortwave Radiation - The photosynthetically active portion (PAR) of downwelling shortwave radiation drives photosynthesis. Because the amount of PAR is approximately proportional to the total amount of shortwave radiation, this

dataset was not separated into the PAR and near infrared portions. Downwelling shortwave radiation data over land were obtained for 1997-2001 from the National Centers for Environmental Prediction (NCEP) reanalysis (Kalnay et al., 1996).

Surface Air Temperature -- Surface air temperature is positively correlated with the amount of solar radiation reaching primary producers, and hence photosynthesis, as well as with the rates of all metabolic reactions including respiration. Surface air temperature data were obtained from the NCEP/ NCAR Reanalysis Monthly Means (Kalnay et al., 1996).

Precipitation -- Precipitation affects water availability, and therefore enables both plant growth and decay. Droughts can limit both CO_2 uptake and soil respiration, although the impact on the net flux may differ with the length and severity of drought (e.g. Baker et al., 2008). Prolonged droughts also promote forest fires. A precipitation dataset was obtained from the Monitoring Product of the Global Historical Climatology Centre in Germany (Adler et al., 2003).

Palmer Drought Severity Index -- The PDSI tracks atmospheric moisture at the surface of the earth relative to local mean conditions, and is calculated using both precipitation and surface air temperature. The Palmer Drought Severity Index (PDSI) was formulated by *Palmer* (1965) as a hydrological accounting system for the central United States, and was subsequently extended globally by Dai et al. (2004).

Vegetation indices: LAI, NDVI, fPAR -- The Normalized Difference Vegetation Index (NDVI) is the dimensionless normalized difference between solar and infrared surface reflectances. Because leaves absorb solar but reflect infrared radiation, NDVI is a measure of green leafy biomass. Leaf Area Index (LAI) is the total surface area of leaves per unit ground area (m^2/m^2). The absorbed fraction of photosynthetically active radiation (fPAR) is the fraction of incident solar radiation absorbed by plants during photosynthesis. NDVI was sourced from the GIMMS dataset, version g (*Tucker et al.*, 2005) based on radiances from the Advanced Very High Resolution Radiometer (AVHRR). fPAR was estimated from the NDVI data using the average of the simple ratio and NDVI methods (Los et al., 2001; Schaefer et al., 2002; Schaefer et al., 2005),

and LAI was estimated from fPAR by inverting Beer's law assuming leaf radiative characteristics from Sellers et al. (1996b).

Land Cover -- Different land cover types are associated with varying levels of Net Primary Productivity (NPP). The DISCover Global Land Cover dataset, obtained from the Global International Geosphere-Biosphere Program (Loveland et al., 2001), contains 18 categories of land cover derived from satellite imagery recorded from April 1992 through March 1993. This dataset was further binned into six categories: Forest, Shrub, Grassland, Agriculture, Barren (including Urban) and Inland Water, and a percent cover for each of these six land cover categories was calculated at the 3.75° x 5° resolution. Only % Agricultural Land, % Forest Cover, % Shrub Cover, % Grassland and a combined % Forest/ Shrub Cover category were selected for further assessment. These derived land cover variables form a static dataset used for the full study period.

Population Density -- Fossil fuel emissions generally trend well with human population density, although densely-populated but poorer countries (e.g. Bangladesh, which is 9th in the world in population, but 69th in emissions (Marland et al., 2006; Central Intelligence Agency, 2007)) weaken this relationship. The population density dataset used in this study was created by Environment Canada with support from the United Nations Environment Programme (Li, 1996).

GDP Density -- A global gridded Gross Domestic Product (GDP) dataset, representing the total economic output of the population living in a given area, was sourced from the International Satellite Land Surface Climatology Project Initiative II Data Collection (Yetman et al., 2004). The population and GDP datasets are static snapshots of the year 1990, and both are normalized by gridcell area to create a density indicator (people/m^2or $USD/$m^2$).

4.2.2.3 Variable selection using the Variance-Ratio test

For this application, a trend with 12 monthly latitudinal land gradients and 12 monthly ocean constants is set as the initial model $\mathbf{X}_0$ for the Variance-Ratio Test, as described in Section 4.2.3.1 and Chapter 3. The Variance-Ratio Test is then run for

each of the 14 candidate auxiliary variables (Table 4.1), adding each individually into $\mathbf{X}_1$ (i.e. $q = 1$). A single variable that significantly improves the trend is selected for inclusion, and this augmented trend becomes the new $\mathbf{X}_0$. The test is then performed again using each of the other 13 remaining variables. Multiple rounds of the test are performed until no significant variables remain at the $\alpha = 0.05$ significance level. Only a single variable is added in each round, even if more than one variable represents a significant improvement to the model. The choice among significant variables is based on its relative level of significance, as well as the importance of its known association with key flux drivers (i.e. photosynthesis, respiration, fossil fuel emissions, etc.).

4.2.3 Covariance matrices (Q and R)

Covariance parameters for land and ocean fluxes in the flux covariance matrix ($\mathbf{Q}$) are optimized separately, and no correlation is assumed between them, as described in *Mueller et al.* (2008). The model-data mismatch variances in the $\mathbf{R}$ matrix, which include measurement, transport, and representation errors for each observation, are assumed to be proportional to the square of the residual standard deviation (RSD) of flask observations from a smoothed curve (GLOBALVIEW- CO_2, 2008), as used in the TransCom series of studies (e.g. Baker et al., 2006).

The parameters of the $\mathbf{Q}$ and $\mathbf{R}$ matrices are optimized using the Restricted Maximum Likelihood (RML) method (Kitanidis, 1995; Michalak et al., 2004; Mueller et al., 2008), as described in Chapter 3. The covariance parameters for the $\mathbf{Q}$ matrix are optimized using process-based and inventory flux estimates from the Carnegie-Ames-Stanford Approach (CASA) model (Randerson et al., 1997) for monthly net ecosystem production (NEP), Takahashi et al. (2002) for monthly net oceanic carbon exchange, and *Brenkert* (1998) for yearly-averaged fossil fuel and cement production emissions. The scaling parameter, c, applied to the squared RSD's in the $\mathbf{R}$ matrix, is optimized using the atmospheric concentration measurements. Covariance parameters are optimized using the model of the trend ($\mathbf{X}$) derived from the variable

selection process described in Section 4.2.3.3.

4.3 Results and Discussion

This section presents CO_2 fluxes estimated using geostatistical inverse modeling, informed both by atmospheric CO_2 measurements and selected auxiliary environmental data. Results are also compared to those obtained by Mueller et al. (2008) using only the atmospheric data constraint.

4.3.1 Variance Ratio Test and selection of auxiliary variables

The Variance Ratio Test is applied as described in Section 4.2.3.3 to select a subset of auxiliary variables that capture a portion of the flux variability, as inferred using the atmospheric CO_2 observations. As previously mentioned, the approach is complemented with scientific understanding regarding the variables and their relationship to flux processes to select among variables that are significant in each round of the test. Fully automatic model-building procedures are not recommended as a means for identifying the best interpretable model, because such procedures can potentially select models that represent only spurious relationships and can fail when applied to comparable datasets (Judd and McLelland, 1989). Note that the Variance Ratio Test determines the significance of the linear relationship between surface flux and auxiliary variables as identified through the relatively sparse atmospheric measurement network. Therefore, selected variables may be more representative of relationships in well-constrained regions.

GDP Density is selected in the first round of auxiliary variable selection (Table 4.1) because it significantly improves the trend, and is believed to best isolate the fossil fuel emission signal, which is the largest single net source of CO_2 on annual timescales. Leaf Area Index is selected in the second round for its association with NPP, and because it is the most significant among the three vegetation indices. For all subsequent rounds, the most significant variable is selected for inclusion in the augmented model of the trend. These variables are fPAR, % Shrub Cover, and Population Density, in the

third, fourth, and fifth rounds, respectively. No additional variables are significant beyond the fifth round. Results from the Variance-Ratio Test also confirm that the monthly latitudinal gradients are a significant improvement upon the monthly land constants implemented in Mueller et al. (2008), a result which holds regardless of whether or not auxiliary variables are also included in the analysis.

Table 4.1: Auxiliary variables and their observed significance levels for each round of the Variance Ratio Test. Variables included in the model of the trend are in bold and their significance levels in the final round before inclusion are highlighted.

Variable	*Round 1*	*Round 2*	*Round 3*	*Round 4*	*Round 5*	*Round 6*
GDP Density	**$<10^{-16}$**					
Population Density	$<10^{-16}$	$3x10^{-8}$	$3x10^{-10}$	$2x10^{-8}$	**$3x10^{-5}$**	
LAI	0.37	**$3x10^{-3}$**				
fPAR	$7x10^{-6}$	0.39	**$<10^{-16}$**			
NDVI	$4x10^{-4}$	0.80	$<10^{-16}$	0.31	0.04	0.41
Shortwave Radiation	0.03	0.29	$5x10^{-5}$	0.64	0.01	0.12
Surface Air Temperature	10^{-4}	0.02	$4x10^{-6}$	$2x10^{-3}$	$3x10^{-3}$	0.06
Precipitation	$6x10^{-9}$	$7x10^{-4}$	$1x10^{-11}$	$3x10^{-3}$	0.02	0.25
PDSI	0.84	0.72	0.55	0.44	0.51	0.77
% Agricultural Land	$<10^{-16}$	10^{-7}	$6x10^{-14}$	$3x10^{-4}$	0.06	0.81
% Forest Cover	10^{-8}	$3x10^{-5}$	$<10^{-16}$	0.03	0.97	0.51
% Forest/ Shrub Cover	0.87	0.59	0.04	$3x10^{-7}$	0.97	0.51
% Grassland	0.12	0.30	0.81	0.01	0.22	0.38
% Shrub Cover	$3x10^{-9}$	$2x10^{-7}$	$6x10^{-6}$	**$2x10^{-11}$**		

Overall, the selected variables are associated with different drivers of terrestrial CO_2 flux, including photosynthesis, respiration, land cover type and fossil fuel emissions. Additional auxiliary variables and/ or functional forms could be applied in the future, in order to capture additional processes (e.g. biomass burning, deforestation and oceanic productivity/ gas exchange) and identify more complex or regional relationships

between auxiliary variables and CO_2 flux variability. However, given that the geostatistical inversion estimates both the model of the trend and flux deviations from this trend, any processes that are not represented by the auxiliary variables can still be represented in the final best estimates of flux, as part of the stochastic component of the best estimate.

4.3.2 Optimized covariance parameters

The optimized parameters for the covariance matrices ($\mathbf{Q}$ and $\mathbf{R}$) are presented in Table 4.2 for the model of the trend presented in the last section, as well as the setup implemented in Mueller et al. (2008). Both land $\mathbf{Q}$ parameters (σ_Q^2 and l_Q) show a significant decrease of approximately 30% from the simple to the complex trend. The optimized scaling parameter (c) for $\mathbf{R}$ decreases by 8%, a smaller but also significant change. Given the absence of any oceanic variables in the complex trend, the ocean $\mathbf{Q}$ parameters remain unchanged between the two trends.

Table 4.2: Optimized model-data mismatch ($\boldsymbol{R}$) and spatial covariance ($\boldsymbol{Q}$) parameters with +/- 1 standard deviation for simple (Mueller et al., 2008) and complex models of the trend.

	$\boldsymbol{Q_{land}}$		$\boldsymbol{Q_{ocean}}$		$\boldsymbol{R}$
Trend	σ^2, $(\mu mol_{CO2}/(m^2 s))^2$	*l,* km	σ^2, $(\mu mol_{CO2}/(m^2 s))^2$	*l,* km	*c*
Simple	0.40 +/- 0.03	2700 +/- 200	0.0030 +/- 0.0003	5700 +/- 500	0.63 +/- 0.04
Complex	0.28 +/- 0.01**	1800 +/- 100**			0.58 +/- 0.04*

*One standard deviation reduction from simple to complex trend

**Two standard deviation reduction from simple to complex trend

The reduction in the model-data mismatch parameter (c) and the land $\mathbf{Q}$ variance parameter (σ_Q^2) provide additional confirmation that the complex trend represents a better representation of the spatial variability of CO_2 flux relative to the simple trend. The reduction in the estimated model-data mismatch demonstrates that fluxes estimated using the complex trend are better able to reproduce the atmospheric concentration measurements relative to those derived using the simple trend. The decrease in σ_Q^2 indicates that, as more of the flux variability is explained by an improved trend, the flux residuals decrease in magnitude. In other words, the complex trend explains a larger fraction of the inferred variability of CO_2 fluxes. Shorter correlation lengths in the residuals also indicate that more of the large scale spatial variability is being captured by the complex model of the trend, leading to residuals that are correlated on smaller scales. As will be discussed in Section 4.3.5, the changes in the $\mathbf{Q}$ and $\mathbf{R}$ parameters also lead to a decrease in grid-scale *a posteriori* uncertainties for the best estimates of flux.

4.3.3 Estimated drift coefficients $(\widehat{\boldsymbol{\beta}})$ and contributions to CO_2 flux $(X\widehat{\boldsymbol{\beta}})$

The estimated drift coefficients ($\widehat{\boldsymbol{\beta}}$) corresponding to the auxiliary variables, their coefficients of variation ($\boldsymbol{\sigma}_{\widehat{\beta}}/\widehat{\boldsymbol{\beta}}$), and the correlation coefficients (ρ) among them are presented in Table 4.3. A positive sign on the drift coefficients indicates a positive correlation with CO_2 flux (i.e. a source or reduction in sink), while a negative sign indicates a negative correlation (i.e. a sink or reduction in source). A coefficient of variation less than 0.5 implies a significant contribution to the trend at the $2\sigma_{\widehat{\beta}}$ level, and all drift coefficients on the auxiliary variables are therefore significant at the 95% level.

The recovered signs on the drift coefficients for the five auxiliary variables show that the inversion is able to infer reasonable relationships between these parameters and CO_2 flux. GDP and Population Densities are associated with sources, as expected given their correlation with fossil fuel emissions, while the opposite signs on LAI and fPAR imply that these variables collectively represent the opposing photosynthesis and

respiration signals. These results lend support to the validity of the Variance Ratio Test for selecting auxiliary variables, as well as provide indirect evidence that the improved model of the trend is able to correctly represent flux variability in the final flux estimates, particularly in under-constrained regions.

Table 4.3: Estimated drift coefficients ($\widehat{\boldsymbol{\beta}}$), coefficients of variation ($\boldsymbol{\sigma}_{\widehat{\boldsymbol{\beta}}}/\widehat{\boldsymbol{\beta}}$), annual average global contribution to flux ($\boldsymbol{X\widehat{\beta}}$) and correlation coefficients ($\boldsymbol{\rho}$) between auxiliary variables in the model of the trend. Also shown is the range of monthly values for the individual $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\sigma}_{\widehat{\boldsymbol{\beta}}}/\widehat{\boldsymbol{\beta}}$ for the land latitudinal gradients and ocean averages, as well as their annual average global contribution to flux. The annual average contribution to flux of the complete trend represents a sum of the contributions by each of the previous components.

	$\widehat{\boldsymbol{\beta}}$[a]	$\boldsymbol{\sigma}_{\widehat{\boldsymbol{\beta}}}/\widehat{\boldsymbol{\beta}}$	$\mathbf{X\widehat{\beta}}$ (GtC/yr)	$\boldsymbol{\rho}$ GDP	Pop	LAI	fPAR	Shrub
GDP Density (000's \$/(m^{2}yr))	180	0.35	1.6	1.00	---	---	---	---
Population Density (people/m^2)	1700	0.26	3.2	-0.42	1.00	---	---	---
LAI (m^2/m^2)	-0.49	0.08	-42.1	0.02	0.01	1.00	---	---
fPAR (*unitless*)	1.9	0.08	46.8	-0.12	-0.12	-0.94	1.00	---
% Shrub Cover (%)	-0.0038	0.18	-3.5	0.14	0.23	0.06	-0.21	1.00
Land Latitudinal Gradients	-0.5 to 0.3	0.2 to 1.1	0.6					
Ocean Constants	-0.08 to 0.01	0.2 to 6.7	-2.8					
Complete Trend			3.8					

[a] *The drift coefficients ($\widehat{\boldsymbol{\beta}}$) have units of $\mu mol_{CO2}/(m^2 s)$ divided by the units of the individual auxiliary variables. Due to differences in units on the auxiliary variables, the magnitudes of the drift coefficients are not directly comparable.*

The annually-averaged global contribution to flux ($\mathbf{X}_i\widehat{\boldsymbol{\beta}}_i$) in GtC/yr is also displayed in Table 4.3 for each of the auxiliary variables, which makes it possible to assess the magnitudes of the recovered drift coefficients in consistent units. GDP and Population Densities together contribute 4.8 GtC/yr globally, which is approximately 70% of the estimated 6.7 GtC/yr global source from fossil fuels and cement production

over this period (Baker et al., 2006; Energy Information Administration, 2002; Marland et al., 2006).

LAI and fPAR have the largest annually-averaged contributions to flux among the different components of the trend. These datasets have similar spatial patterns, and this collinearity implies that the interpretation of their combined contribution to flux is more reliable than their individual relationships to flux, as demonstrated by the strong anticorrelation between their estimated drift coefficients ($\rho = -0.94$). The combined contribution of LAI and fPAR within the trend shows net sources and sinks on a seasonal basis that are consistent with the expected biospheric signal. This contribution also plays a large role in defining the spatial variability of the overall terrestrial flux estimates (as shown in Figure 4.1). The combined annually-averaged global contribution to flux of LAI, fPAR and % Shrub is 1.2 GtC/yr, implying that these variables together represent a large portion of the biospheric signal, which has a strong seasonality but a relatively small annually-averaged net flux.

The positive drift coefficient associated with fPAR (representing sources or reductions in sinks) and the negative drift coefficient associated with LAI (representing the opposite) appear to contradict process-based understanding of the relationship between these variables and biospheric CO_2 fluxes. Photosynthesis is frequently estimated from fPAR, given assumed rates of autotrophic respiration (*Tucker and Sellers*, 1986; *Potter et al.*, 1993), while LAI, as a measure of biomass, is more commonly associated with autotrophic and heterotrophic respiration (e.g. *Reichstein et al.*, 2003). However, at the spatial and temporal resolution of this study, LAI appears to capture the strong seasonality expected for photosynthesis, while the weaker seasonal cycle of fPAR captures variability expected for total ecosystem respiration (Figure 4.2).

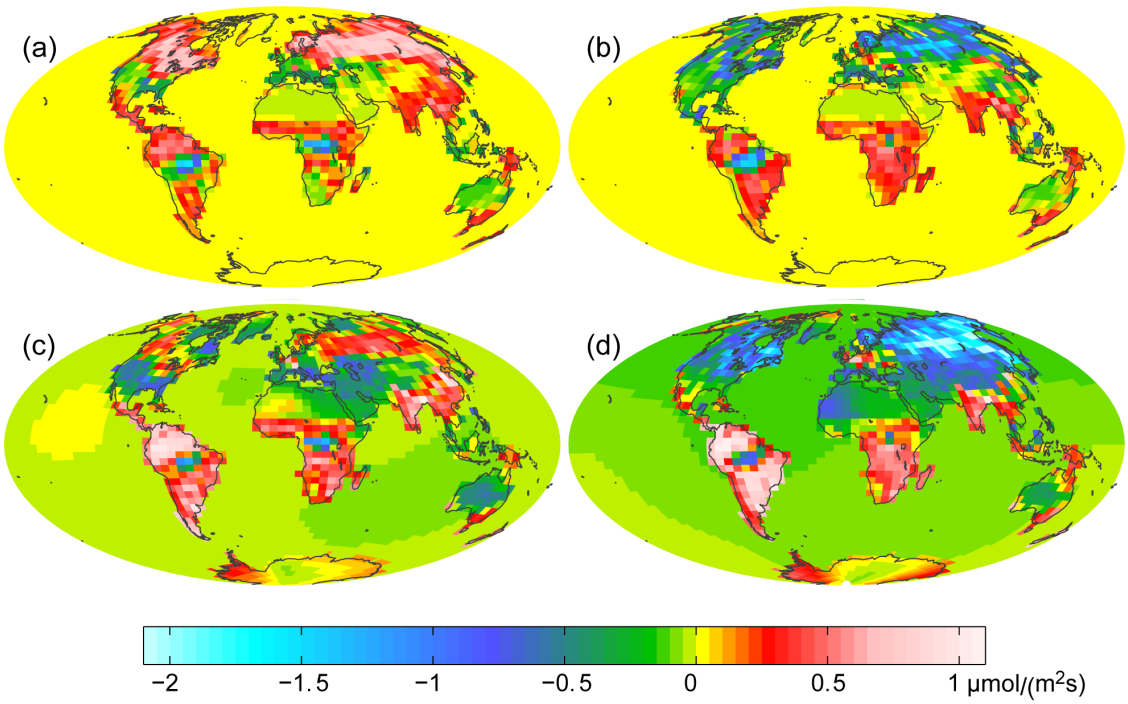

Figure 4.1 (a) Contribution to flux estimates by LAI and fPAR within the model of the trend ($\boldsymbol{X\hat{\beta}}$) for May 2000, (b) contribution by LAI and fPAR for July 2000, (c) best estimates of flux ($\hat{\boldsymbol{s}}$) for May 2000, and (d) best estimates of flux ($\hat{\boldsymbol{s}}$) for July 2000.

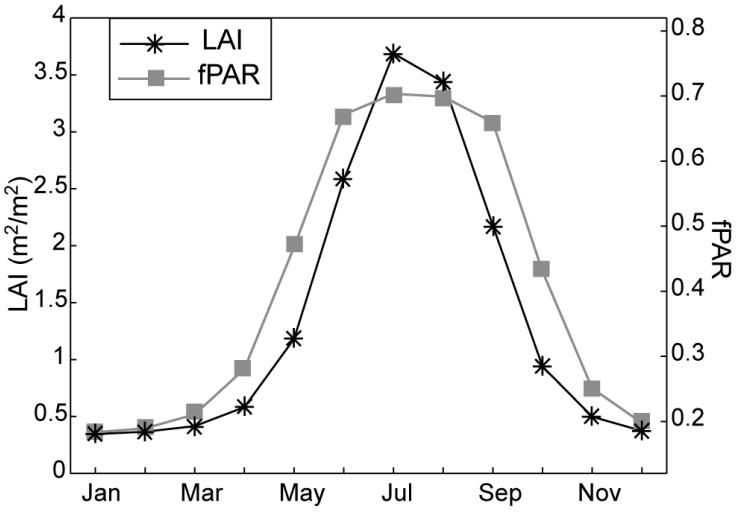

Figure 4.2: Average monthly LAI and fPAR (from 1997 to 2001) for the combined Northern Hemisphere land regions of Boreal Asia, Europe, and Boreal North America (as defined in Figure 4.7).

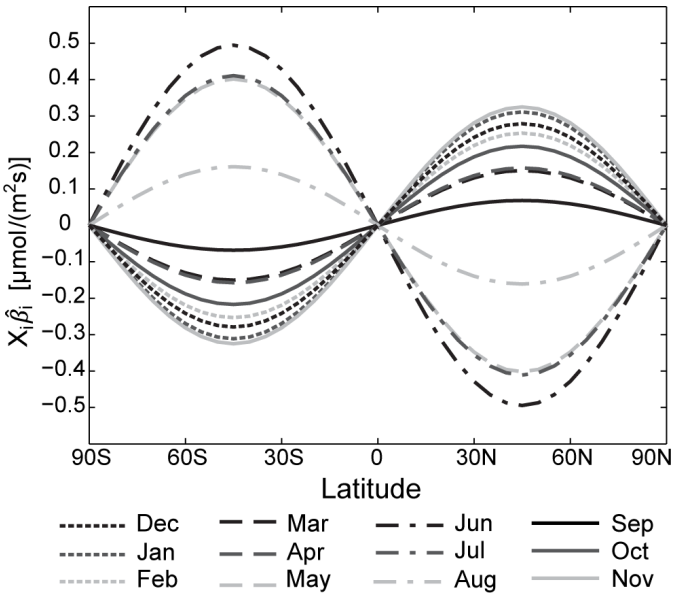

Figure 4.3: Contribution to flux by the 12 monthly latitudinal land gradients within the model of the trend ($\boldsymbol{X\hat{\beta}}$).

Figure 4.3 shows the contribution to the trend ($\mathbf{X}_i\hat{\boldsymbol{\beta}}_i$) by the monthly terrestrial latitudinal gradients, which show strong seasonality. For example, the latitudinal gradient in June shows a sink in the Northern Hemisphere mid-latitudes with a corresponding source in the Southern Hemisphere, while the gradient shows the opposite flux pattern in January. This result demonstrates that the atmospheric data are able to correctly identify seasonal variability between the hemispheres that is unexplained by the other auxiliary variables within the trend. Eight of the twelve multipliers show a source in the Northern Hemisphere, likely as a result of the year-round fossil fuel CO_2 sources from industrialized areas in North America, Europe and Asia that are not captured by the contributions of GDP Density and Population Density within the model of the trend.

The complete model of the trend including the latitudinal gradients, ocean constants and auxiliary variables, represents a 3.8 GtC/yr annually-averaged source to the atmosphere from 1997 to 2001. The overall annually-averaged global flux estimate from the inversion is a source of 4.0 GtC/yr, which indicates that the complex

model of the trend captures approximately 95% of the global atmospheric increase on an annually-averaged basis, and is therefore explaining a substantial portion of total flux at this aggregated scale. As shown in equation 3.8, the residual component of the flux estimates are explained by the stochastic component, such that the additional source apparent in the atmospheric measurements but not captured by the trend is still incorporated into the final flux estimates.

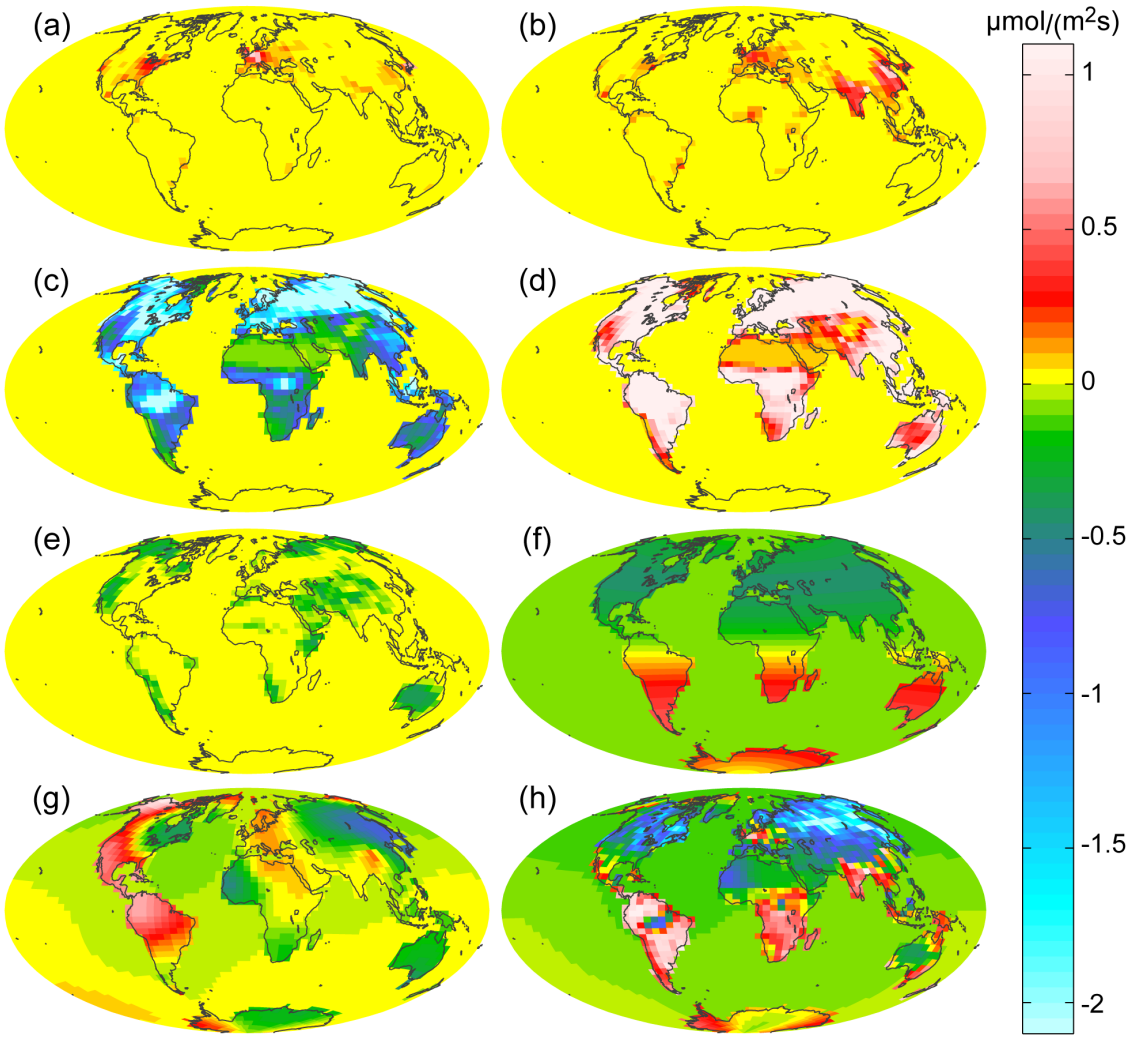

Figure 4.4: Contribution of various components within the model of the trend ($\mathbf{X}\hat{\boldsymbol{\beta}}$) toward the best estimates of flux ($\hat{\mathbf{s}}$) in July 2000: (a) GDP density, (b) population density, (c) LAI, (d) fPAR, (e) % Shrub Cover, (f) latitudinal gradient and ocean constant, (g) stochastic component of best estimates, and (h) full best estimates ($\hat{s}$).

4.3.4 Spatial distribution of the deterministic and stochastic components of a posteriori flux estimates $(\hat{\boldsymbol{s}})$

Figure 4.4 illustrates the spatial distribution of each component of the model of the trend ($\mathbf{X}\hat{\boldsymbol{\beta}}$), the spatially correlated flux residuals ($\mathbf{Q}\mathbf{H}^T\boldsymbol{\Psi}^{-1}\left(\mathbf{z}-\mathbf{H}\mathbf{X}\hat{\boldsymbol{\beta}}\right)$), and the best estimates of flux $(\hat{\mathbf{s}})$ for July 2000. The shrublands in arid regions like Australia and the boreal regions of North America and Asia show small negative contributions to the overall flux, while LAI and fPAR show large, but opposite, contributions to flux in vegetated areas, as previously discussed in Section 4.3.3. Both GDP and Population Densities show positive contributions to flux, although their spatial patterns differ. The terrestrial latitudinal flux gradient reflects climatic variability unexplained by the other auxiliary variables, and shows the largest negative contribution to flux in the Northern Hemisphere mid-latitudes for this month. It should be noted that % Shrub Cover, GDP Density and Population Density are static datasets and therefore, the July 2000 contributions of these variables shown in Figure 4.4 represent only long-term average contributions to flux.

While the magnitude of the stochastic component is generally reduced as the ability of the trend to explain flux variability becomes stronger (as evidenced by the reduction in the σ_Q^2 land parameter shown in Table 4.2), the stochastic component associated with the flux estimates in July 2000 is still responsible for positive contributions over South America and most of North America, and slight negative contributions in northeast Asia, Australia and parts of Africa. In fact, the stochastic component adds a positive contribution to flux in tropical Central and South America for approximately eight months of each year of the inversion. This shows that although the complex model of the trend cannot capture a systematic flux signal in this region, possibly due to the lack of auxiliary variables associated with biomass burning and/or deforestation, the stochastic component identifies a net additional source in these regions.

In the Mueller et al. (2008) study, $\mathbf{X}\hat{\boldsymbol{\beta}}$ is simply an average flux over land and an average flux over oceans for each calendar month. Therefore, the spatial variability of

the best estimates at the grid-scale is entirely determined by the spatially-correlated stochastic component. In contrast, for the complex trend with auxiliary variables, each component within the trend adds an additional layer of spatial variability to the *a posteriori* flux estimates, weighted by that component's estimated relationship to flux ($\widehat{\boldsymbol{\beta}}$). Therefore, the complex trend inversion is able to more realistically represent grid-scale variability without relying on the use of explicit prior flux estimates used in synthesis Bayesian inversions.

4.3.5 A posteriori grid-scale uncertainty reduction from simple to complex trend

The greater ability of the complex trend to capture flux variability relative to the simple trend, represented by a reduction in the optimized land variance parameter in the $\mathbf{Q}$ matrix and the scaling parameter in the $\mathbf{R}$ matrix, leads to an overall decrease in *a posteriori* uncertainty on the flux estimates (see equation 3.5). Figure 4.5 shows the average percent change in uncertainty at the gridscale between the simple and the complex trend inversion for the year 2000. The uncertainty on land is reduced by up to 14%, with higher decreases in areas such as Africa, South America and Southeast Asia. This is due to the fact that these areas, under-constrained by the atmospheric measurements, are now informed by a better deterministic model of the trend. For the oceans, the uncertainty is reduced by approximately 2% for most regions.

Whereas the reduction in the variances in $\mathbf{Q}$ and $\mathbf{R}$ lead to a decrease in the *a posteriori* uncertainties, including additional variables in the model of the trend also leads to additional uncertainties resulting from the estimation of the corresponding drift coefficients ($\widehat{\boldsymbol{\beta}}$). The uncertainty associated with estimating $\widehat{\boldsymbol{\beta}}$ contributes to the *a posteriori* uncertainties through the term $-\mathbf{XM}$ in equation 3.5 (this term is always positive). Therefore, some regions actually show a slight increase in the *a posteriori* uncertainty when moving from the simple to the complex trend. For example, the high values of GDP and Population Densities in the Northeastern United States, Germany, China, Japan and Bangladesh, lead to increases in estimated uncertainty of up

to 11%.

The general reduction in *a posteriori* grid-scale uncertainty ($\sigma_{\hat{s}}$) from the simple to the complex trend shown in Figure 4.5 leads to a small increase in the number of significant terrestrial sources and sinks estimated at the grid scale (17% of gridcells for the simple trend vs. 25% for the complex trend at the $1\sigma_{\hat{s}}$ level, or 2% vs. 6% at the $2\sigma_{\hat{s}}$ level). Overall, however, grid-scale uncertainties are high relative to flux magnitudes in both inversions due to the limited network of atmospheric measurements, as expected. Note that the reduction in uncertainty from the simple to the complex trend described here is not analogous to the reduction in uncertainty described in synthesis Bayesian inversion studies (e.g. Rödenbeck et al., 2003; Baker et al., 2006). In synthesis Bayesian inversions, the *a priori* uncertainty is described by the matrix $\mathbf{Q}$, whereas the *a priori* uncertainty in geostatistical inversions is effectively infinite given that there are no *a priori* assumptions about the drift coefficients $\boldsymbol{\beta}$. Instead, the reduction in uncertainty reported here represents the relative constraints on fluxes achieved by two different inversion setups, namely those described by the simple and complex trends.

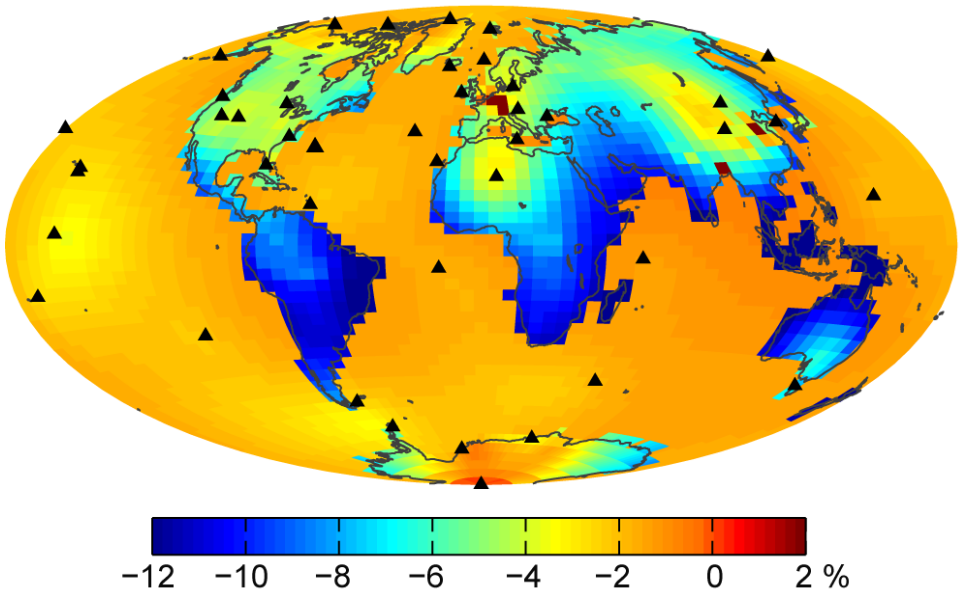

Figure 4.5: Percent change in a posteriori uncertainty ($\sigma_{\hat{s}}$) from the simple to the complex trend inversion, annually averaged for year 2000. Triangles represent measurement locations.

4.3.6 Continental-scale seasonal cycle for year 2000

Figure 4.7 presents monthly flux estimates and $1\sigma_{\hat{s}}$ confidence intervals for the year 2000 resulting from the simple and complex trend inversions, aggregated to the 22 TransCom regions (e.g. Gurney et al., 2003) shown in Figure 4.6. In some regions, such as Boreal North America, Temperate North America and Northern Africa, results from the application of the two trends are nearly identical. In other regions, the auxiliary variables and terrestrial latitudinal gradient in the complex trend have an impact on the flux estimates. For example, the complex trend inversion shows a larger summertime sink in Boreal Asia and Europe and a slightly higher year-round flux in Tropical Asia, with this latter result most likely due to the positive contribution to flux associated with densely-populated areas in Bangladesh and southern China. The better constraint on terrestrial fluxes provided by the improved trend also slightly alters fluxes in nearby ocean regions.

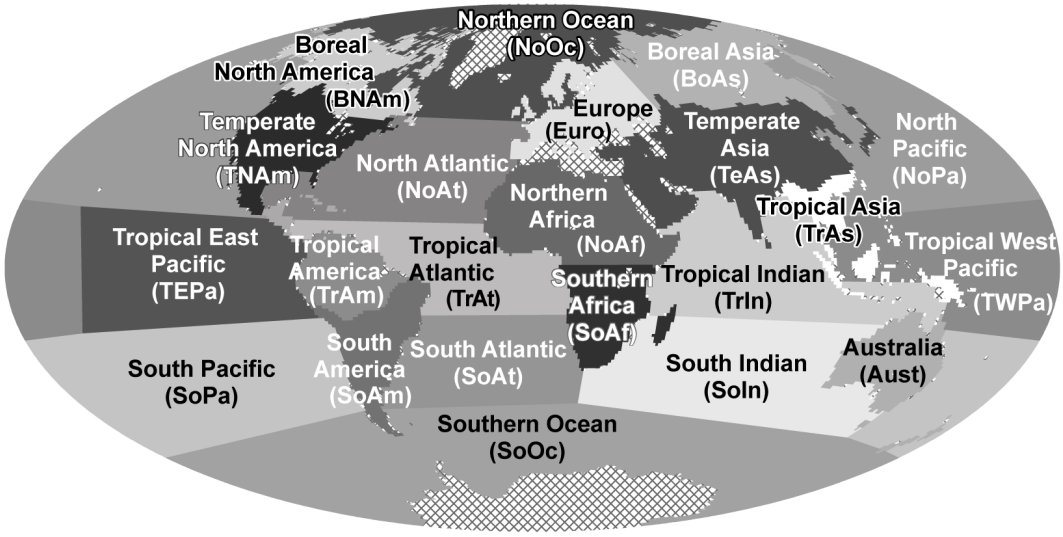

Figure 4.6: Locations of 11 land and 11 ocean TransCom regions (e.g., Gurney et al., 2003).

However, apart from these small differences, the magnitude and seasonality of aggregated fluxes inferred using the two trends agree well for both land and ocean regions. This result shows that there exists a relatively strong atmospheric constraint on the seasonal cycle of geostatistical flux estimates at the scale of the 22 TransCom regions. This is important, given that flux patterns at the grid scale vary significantly between the two inversions, and supports the hypothesis that the flux estimates at the aggregated scale are representative of the information content of the atmospheric data.

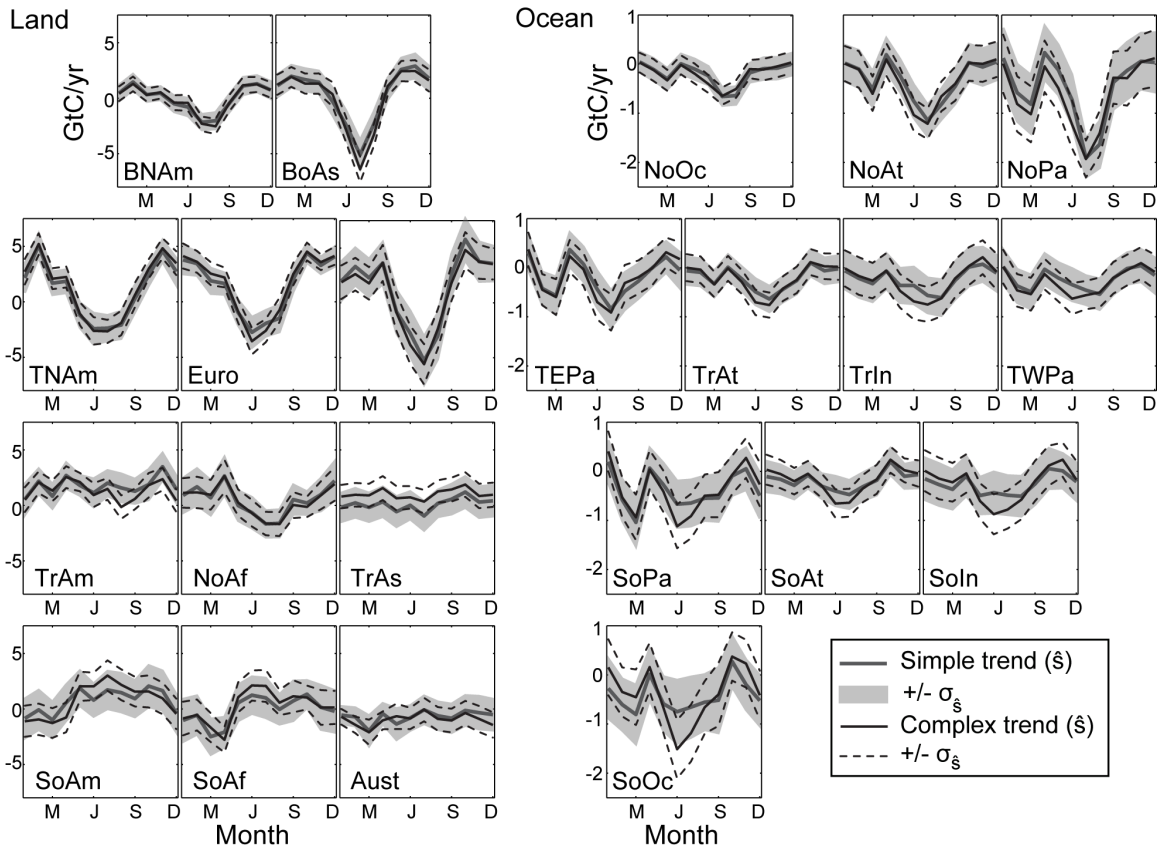

*Figure 4.7: Monthly best estimates (**ŝ**) aggregated to 22 TransCom regions with 1$\sigma_{\hat{s}}$ confidence intervals for year 2000 for simple (Mueller et al., 2008) and complex trend inversions.*

4.3.7 Annually-averaged continental-scale sources and sinks

Figure 4.8 presents annually-averaged fluxes for 1997 to 2001 from the simple and complex trend inversions, aggregated to the 22 TransCom regions. Uncertainty associated with the annually-averaged fluxes is 7% to 19% lower for land regions and 2% to 7% lower for ocean regions in the complex trend inversion relative to the simple trend inversion, demonstrating that the improved trend helps to better constrain flux estimates at aggregated spatial and temporal scales, as well as at the grid-scale (as discussed in Section 4.3.5).

For the complex trend, most land regions show significant ($1\sigma_{\hat{s}}$) net sources, whereas Boreal North America and Boreal Asia are flux-neutral, and Australia is a significant sink. The predominance of continental-scale terrestrial sources reflects the impact of fossil fuel emissions on the annually-averaged CO_2 fluxes. An analysis of the biospheric annually-averaged flux, derived by subtracting fossil fuel inventory data (*Brenkert*, 1998) from the annual total values shown in Figure 4.8, shows that Temperate North America, Europe, Temperate Asia and Australia all act as significant biospheric sinks ($1\sigma_{\hat{s}}$) in the complex trend inversion.

For all ocean regions, fluxes from both inversions show a significant ($1\sigma_{\hat{s}}$) sink, and the results from the two inversions are not significantly different from one another. However, as discussed in Mueller et al. (2008), the relatively constant oceanic flux estimates across regions reflect the limited information content of the atmospheric measurements, with oceanic flux estimates in many regions remaining close to the global averages reflected in the model of the trend. Despite the lack of oceanic auxiliary variables, a better constraint on terrestrial fluxes within the complex trend reduces the strength of the overall ocean sink (from -3.0 GtC/yr to -2.7 GtC/yr), bringing these estimates into closer agreement with independent results from extrapolated ocean ship-track data (Takahashi et al., 2002) and inverse modeling studies that make direct use of these bottom-up estimates (Rödenbeck et al., 2003; Baker et al., 2006).

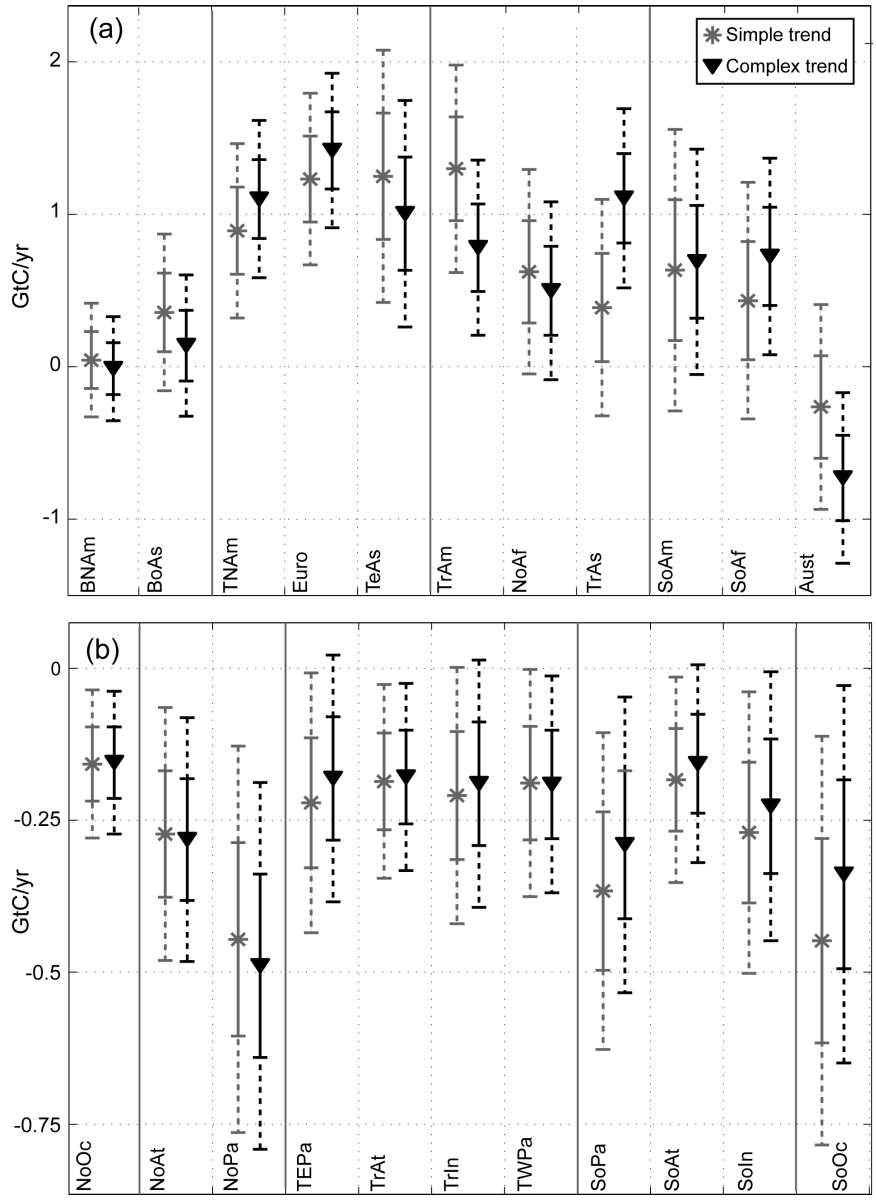

Figure 4.8: Annually averaged flux for simple and complex trend inversions for TransCom (a) land and (b) ocean regions for 1997 to 2001. Land fluxes include both biospheric and fossil fuel components. Error bars represent $1\sigma_{\hat{s}}$ and $2\sigma_{\hat{s}}$ confidence intervals.

A few under-constrained land regions, such as Tropical Asia, Tropical America and Australia, show significant ($1\sigma_{\hat{s}}$) changes in estimated average flux between the two inversions. The significant increase in Tropical Asia and decrease in Tropical America demonstrate that the addition of auxiliary information with global coverage helps to

constrain regions remote from measurement locations (see map in Figure 4.6). In fact, the estimates obtained using the complex trend are closer to bottom-up estimates for these regions. For example, CASA estimates of net ecosystem exchange (NEE) (Randerson et al., 1997) with regional corrections for deforestation and re-growth, as applied in Baker et al. (2006), and fossil fuel emission estimates from Brenkert (1998) yield a 1.4 GtC/yr source for Tropical Asia and a 0.7 GtC/yr source for Tropical America, which are similar to the independent estimates obtained using the complex trend inversion. The significant decrease in the net flux from Australia, however, is not consistent with estimates from previous inverse modeling studies (Rödenbeck et al., 2003; Baker et al., 2006) and bottom-up models, which show a near-neutral biospheric flux. The stronger estimated sink in Australia is likely caused by the negative drift coefficient on % Shrub Cover in the complex model of the trend, along with the large areas of open shrublands in this region. Given that this drift coefficient represents a globally averaged estimated relationship between % Shrub Cover and CO_2 flux, estimates in Australia may be unduly influenced by the relationship between shrublands and flux in the better-constrained boreal regions.

The main conclusion to be drawn from the comparison between the annually-averaged, continental-scale fluxes for the two trends is that, as with the seasonal cycle of continental-scale fluxes, there is a relatively strong atmospheric constraint on fluxes at this scale. This is seen through the consistent results obtained using the simple and complex trend inversions. However, at this temporally-aggregated scale, auxiliary variables can significantly impact the annually-averaged flux estimates for certain under-constrained regions, in a manner consistent with process-based understanding of CO_2 flux. This improvement, however, is contingent on the validity of assuming a global relationship between auxiliary variables and CO_2 flux. Overall, as evidenced by lower *a posteriori* uncertainties, the complex trend inversion is better able to constrain annually-averaged continental-scale fluxes, relative to the simple trend inversion.

4.4 Conclusions

This paper presents a method for incorporating auxiliary information provided by spatially-distributed datasets associated with CO_2 flux processes into a geostatistical inverse modeling approach. This approach is then used to estimate monthly-averaged, global, grid-scale CO_2 fluxes using concentration measurements from a subset of the NOAA-ESRL Cooperative Air Sampling Network. The auxiliary datasets with spatially and temporally heterogeneous global coverage help to constrain flux estimates, especially in regions far from measurement locations, and also help to recover fine-scale flux variability that cannot be inferred through the concentration data alone, due to atmospheric transport and mixing. The resulting flux estimates are more realistic, and have lower uncertainty, than those presented in the Mueller et al. (2008) geostatistical inversion study, which relies only on the information content of the atmospheric data. This conclusion is supported by the physically reasonable relationships ($\widehat{\boldsymbol{\beta}}$) between the auxiliary variables and flux recovered by the inversion, as well as the reduction in grid-scale *a posteriori* uncertainty achieved by the complex model of the trend. The uncertainties on the drift coefficients in the model of the trend are also estimated by the inversion and then incorporated into the final *a posteriori* uncertainties on the flux estimates. This approach differs from synthesis Bayesian inversions, where prior flux magnitudes and uncertainties are pre-specified, and are used directly in estimating *a posteriori* uncertainties. Therefore, the *a posteriori* uncertainties on the flux estimates are more strongly data-driven than those estimated in previous inversion studies. One aspect that will be the subject of ongoing work is the impact of the assumption of a constant global relationship between the auxiliary variables and flux within the model of the trend, which is more strongly affected by fluxes in well-constrained regions.

The Variance Ratio Test is used to determine the combination of the candidate auxiliary variables that is best able to explain the flux variability evident in the atmospheric measurement data. From an initial superset of 14 auxiliary variables, five variables were found to significantly improve the model of the trend. These variables are associated with either biospheric activity or fossil fuel emissions. An analysis of the

estimated drift coefficients on the auxiliary variables shows that LAI and fPAR capture a substantial portion of the combined signal of photosynthesis and respiration. The negative drift coefficient for LAI and the positive one for fPAR are opposite to current understanding of the mechanistic relationship between these variables and CO_2 flux. An analysis of these datasets shows that the weaker seasonality in the fPAR dataset relative to LAI allows this variable to more strongly explain the signal associated with total ecosystem respiration at the scales examined in this study. The drift coefficients for the other selected variables indicate that % Shrub Cover explains residual biospheric sinks (or decreases in sources), while GDP and Population Densities explain approximately 70% of the expected global fossil fuel emission signal.

As reflected in the optimized covariance parameters associated with the flux residuals and the model-data mismatch, the model of the trend implemented in this study is able to explain significantly more of the flux variability evident from the atmospheric data relative to a simple model of the trend containing monthly flux averages over land and ocean, as implemented in Mueller et al. (2008). The reduction in the covariance parameters leads to reduced *a posteriori* uncertainties on the flux estimates of up to 14% for the annually-averaged grid-scale fluxes, and up to 19% at the annually-averaged continental scale. This uncertainty reduction is strongest in under-constrained regions in Africa, South America and Southeast Asia.

A comparison of the seasonal cycle of flux estimates at aggregated continental scales shows no significant differences between the simple trend inversion of Mueller et al. (2008) and the complex trend inversion implemented in this study, pointing to a relatively strong atmospheric constraint at this spatial scale. At the annually-averaged continental scale, the auxiliary variables in the complex trend significantly change flux in a few terrestrial regions under-constrained by the measurement network, in a manner consistent with bottom-up understanding of flux in these regions. However, the stronger inferred sink in Australia shows that a global average linear relationship between auxiliary variables and flux may not be representative for some regions or variables. Apart from these few terrestrial regions, the agreement among both the

monthly and annually-averaged fluxes at the continental scale points to a strong atmospheric constraint on flux estimates at spatially-aggregated scales.

Finally, the geostatistical inverse modeling approach presented here provides a method for validating scale-dependent understanding of the relationship between various datasets associated with CO_2 flux processes and actual CO_2 flux variability, as seen through the existing atmospheric monitoring network. In future work, the use of biospheric model output, non-linear, and regional relationships in the model of the trend could help to differentiate among competing hypotheses about processes controlling flux variability, and thereby contribute to process-based understanding of CO_2 flux drivers. This approach will also continue to improve flux estimates, while minimizing *a priori* assumptions inherent to inversion studies. As such, the geostatistical approach provides a unique opportunity for reconciling top-down and bottom-up estimates of CO_2 flux variability at various spatiotemporal scales.

CHAPTER 5

Regional-scale geostatistical inverse modeling of North American CO_2 fluxes: a synthetic data study

5.1 Introduction

CO_2 fluxes cannot be directly measured at regional scales, and have instead been inferred from atmospheric concentration patterns using inverse modeling techniques. While earlier global inversion studies had used atmospheric CO_2 concentration measurements sampled in the free troposphere at remote or high-altitude locations to infer continental-scale CO_2 fluxes (e.g. Gurney et al., 2002; Baker et al., 2006), the recent convergence of several factors has made it feasible to estimate sub-continental scale CO_2 fluxes in a regional inverse modeling framework (e.g. Peylin et al., 2005; Lauvaux et al., 2008). First, continuous ground-based measurements of atmospheric CO_2 taken at several North American and Eurasian sites (e.g. Bakwin et al., 1998; Haszpra, 1999) provide data with high temporal (and, increasingly, high spatial) resolution to constrain carbon fluxes at finer scales. These continuous measurement locations also tend to be sited in continental, low-altitude areas with strong biospheric activity, providing more information about flux variability at sub-continental scales relative to the measurements used in global inversions. Finally, recent advances in regional atmospheric transport modeling and the use of analyzed wind fields with high spatial resolution make it feasible to appropriately take advantage of continuous data from continental locations in regional inversions.

The use of continuous, continental data in grid-scale CO_2 inversions is relatively new, and, therefore, many questions remain as to the optimal approach for taking advantage of these large and highly variable data streams. Synthetic data (a.k.a.

"pseudo-data") experiments are useful in the design of inversions, because they include a set of specified baseline fluxes with which results can be compared, making it easier to diagnose potential biases in inferred fluxes under a number of different scenarios. The interpretation of flux estimates in a synthetic data inversion is also simplified relative to a real data inversion in two important ways. First, synthetic measurements are only influenced by fluxes occurring within the domain of study, and therefore there is no need to specify boundary conditions. Errors in boundary conditions used in regional real-data inversions can bias flux estimates, particularly for short periods (e.g. Peylin et al., 2005; Göckede et al., 2010b). Second, the effect of atmospheric transport model errors can be controlled by using the same transport model to create the synthetic measurements as is used to estimate fluxes in the inversion.

In addition to the simplifications associated with performing a synthetic data inversion, the geostatistical inversion approach (Michalak et al., 2004) furthermore makes it possible to eliminate the impact of the choice of *a priori* flux estimates on inversions. Geostatistical inversions are Bayesian, but do not prescribe a prior estimate of the flux distribution from biospheric models and/or inventories. Also, covariance parameters can potentially be optimized using the atmospheric measurements themselves, further eliminating reliance on bottom-up model output. Therefore, this approach provides a unique opportunity to assess the information content of the available atmospheric measurement data (Mueller et al., 2008), specifically the impact of using continuous, continental measurements in a regional inversion.

Spatial and temporal aggregation errors (Kaminski et al., 2001) are a particular concern when using data collected in high variability areas to estimate fluxes in these same regions, e.g. regional scales within the North American continent. Spatial aggregation errors occur when fixed flux patterns are imposed for large regions (e.g. Law et al., 2002; Peters et al., 2007) because the inversion cannot adjust the flux patterns within the specified region, even though atmospheric observations are sensitive to sub-regional variability. Aggregation errors can also be temporal as well as spatial, and temporal aggregation errors occur when estimating a single flux, or flux

adjustment, over a time interval with significant intra-period variability, e.g. diurnal and synoptic-scale variations. Geostatistical inversions, as well as some other recent inversions, help to address these concerns by estimating fluxes directly on a grid at fine spatial resolutions, thereby minimizing spatial aggregation errors (e.g.; Gourdji et al. 2008; Schuh et al., 2009). Also, fluxes can be estimated directly at short timescales, a topic explored in the current work.

This study uses a series of synthetic data inversion experiments to evaluate a regional geostatistical grid-scale ($1^\circ \times 1^\circ$) inversion for June 2004 over North America, using the nine CO_2 observing towers operational in the United States and Canada at that time. The primary objectives of this study are threefold. First, we investigate the use of available atmospheric measurements to infer reasonable covariance parameters (both flux covariance and model-data mismatch), as well as the impact of these inferred parameters on the estimated fluxes. Secondly, we perform inversions estimating fluxes at three different temporal resolutions ranging from 3-hourly to 4-day averages, in order to evaluate the impact of temporal aggregation errors on inversion results. The third objective of this study is to assess the information content of the limited atmospheric network for 2004, by comparing inversion results to the "true" fluxes at both the grid and aggregated eco-region scales. The effect of random transport model error is additionally explored throughout this study by adding random noise with realistic magnitude to the synthetic measurements, and then observing their impact on the inversion results.

5.2 Inversion Setup

This section describes the setup of each component of the geostatistical inversion objective function (equation 3.1) implemented for the analyses performed in this study. Using the components described below, equations 3.2 to 3.7 are used to estimate *a posteriori* flux estimates ($\hat{\mathbf{s}}$) and uncertainties ($V_{\hat{\mathbf{s}}}$), in addition to the drift coefficients ($\widehat{\boldsymbol{\beta}}$) and uncertainties ($V_{\widehat{\boldsymbol{\beta}}}$) associated with the trend ($\mathbf{X}$). Please review Chapter 3 for a more complete summary of the geostatistical inversion method.

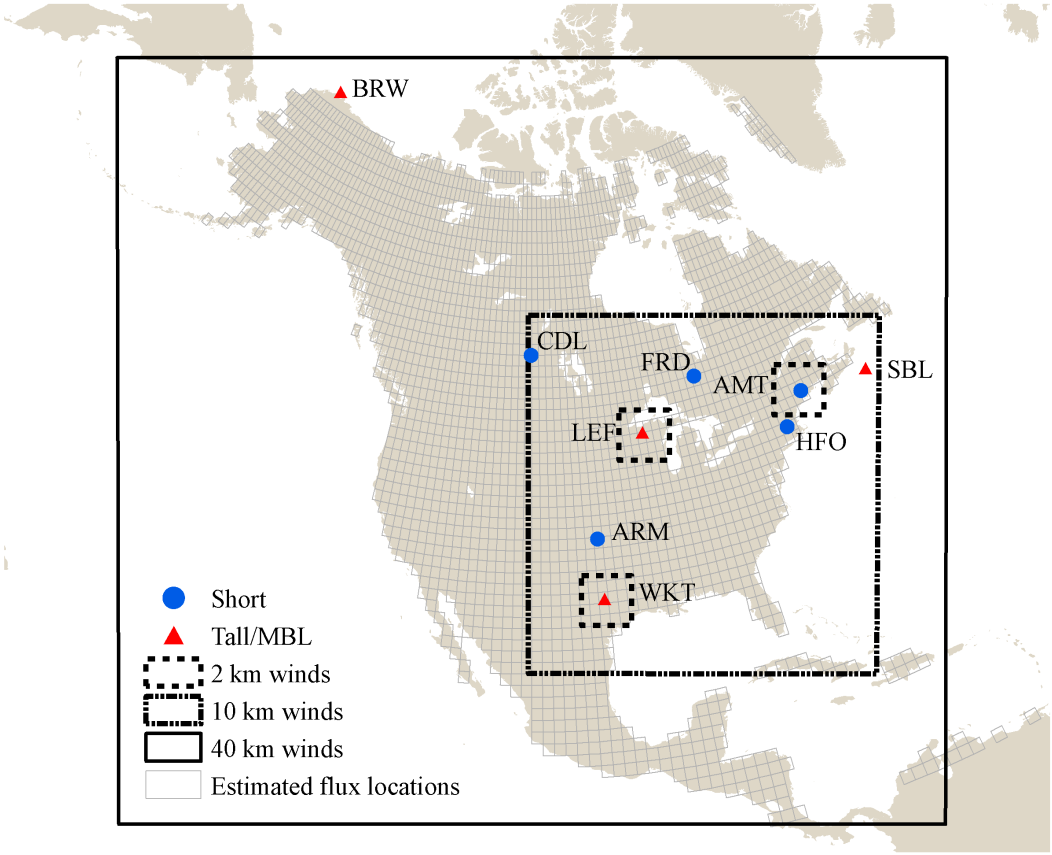

Figure 5.1: Location of nine measurement towers used in the study, as well as the domains for the three levels of high-resolution nesting with the WRF winds. The background grid represents the flux estimation resolution of 1° × 1°.

5.2.1 Flux estimation resolution (s)

For all inversions, fluxes are estimated at a 1°×1° grid-scale spatial resolution, with the domain including all land cells within the range of 10°-70°N and 50°-170°W, yielding 2641 estimation regions (Figure 5.1). Fluxes are estimated from June 1 to July 2, 2004 in universal time (UTC) using three different temporal resolutions: 3-hourly (henceforth referred to as F3hr), a 4-day average diurnal cycle with 3-hourly time increments (F4d-diurnal), and a flat 4-day average without any diurnal variability (F4d). These three temporal resolutions make it possible to investigate the benefit of directly estimating the diurnal cycle of fluxes, and, conversely, the risk of temporal aggregation

error associated with estimating fluxes averaged over multiple days. Despite the potential benefits associated with estimating finer-scale fluxes, the number of estimated fluxes and associated computational costs grow as the temporal resolution becomes finer, as shown in Table 5.1. Additional details associated with the setup for each temporal flux resolution are described in Sections 5.2.3 to 5.2.7.

Table 5.1: Inversion characteristics for the three flux temporal resolutions.

Case	Flux Resolution	# of Estimated Fluxes	Structure of trend (X)	Structure of flux covariance (Q)
F4d	4-day average	$(2641 \times 8) = 21{,}128$	One spatiotemporal mean	Full spatiotemporal flux covariance
F4d-diurnal	4-day average diurnal cycle (with 3-hourly bins)	$(2641 \times 64) = 169{,}024$	Eight spatial means by 3-hourly bins	Full spatiotemporal flux covariance across 4-day periods, but not within diurnal cycle
F3hr	3-hourly	$(2641 \times 256) = 676{,}096$	Eight spatial means by 3-hourly bins	Full spatiotemporal flux covariance across days, but not within diurnal cycle

5.2.2 Atmospheric transport (H)

Atmospheric transport models are necessary for CO_2 inversions in order to quantify the sensitivity of measured concentrations to surface fluxes, or the concentration footprints that populate the atmospheric transport matrix $\mathbf{H}$. The Stochastic Time-Inverted Lagrangian Transport Model (STILT) model (Lin et al., 2003) is used for the current study. STILT, which has already been applied in several pilot studies aimed at constraining CO_2 sources and sinks in the United States (Gerbig et al., 2003, 2006; Lin et al., 2004; Matross et al., 2006), represents air arriving at observation locations as an ensemble of particles that are transported backward in time. The

particle velocities in STILT are in turn derived from meteorological fields generated by gridded numerical weather prediction models, in this case from the Weather Research & Forecasting (WRF) model (Skamarock et al., 2005), version 2.2. For this study, WRF v2.2 was configured to use three levels of high resolution nesting: a 2-km resolution grid around the three tallest measurement towers (LEF, AMT and WKT, see Table 5.2 and Figure 5.1), embedded in a 10-km resolution grid over the northern Midwest, Gulf Region, and New England extending to approximately 105°W, and then an outermost 40-km resolution grid covering the rest of the overall domain of the inversions.

At each measurement location, 10-day back-trajectories of 500 particles were generated using STILT every hour from June 1 to July 8, 2004. Concentration footprints, or sensitivities, are then calculated at 3-hourly intervals back in time, by integrating these particle trajectories over the North American 1°×1° grid as described in Lin et al. (2003). Finally, these high-resolution $\mathbf{H}$ matrices are aggregated to the temporal resolution of the concentration data, described in Section 5.2.4, and the three flux temporal resolutions.

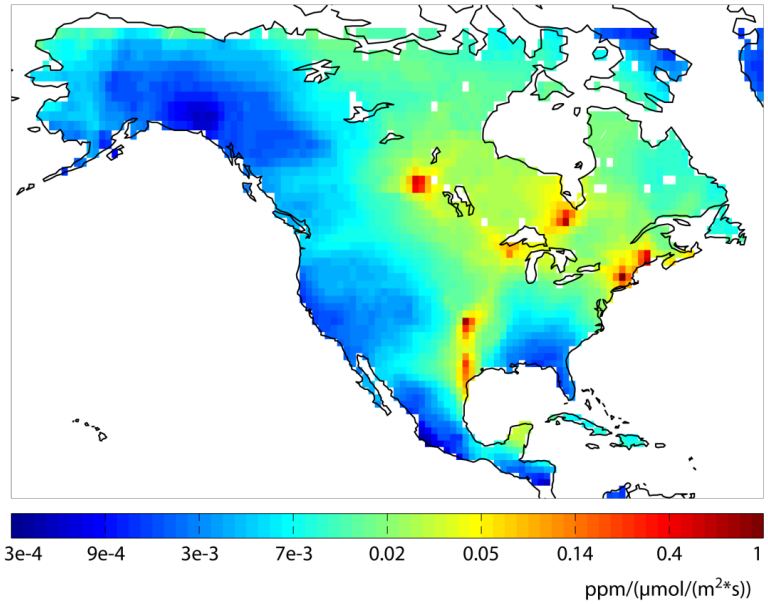

Figure 5.2: Average sensitivity of measurements in June 2004 at nine towers to all prior fluxes.

A map of the average sensitivity of measurements to fluxes for June 2004, derived from the concentration footprints, is shown in Figure 5.2. As seen here, many parts of North America are not well-constrained by the 9-tower measurement network in 2004. These areas include northwest Canada and Alaska, the southwestern and southeastern United States, and parts of Central America. In contrast, the eastern temperate forests and Midwestern agricultural areas have a stronger atmospheric data constraint.

5.2.3 Synthetic concentration time series ($\mathbf{z}$)

One goal of this study is to assess the projected accuracy of North American estimates of CO_2 flux using a contemporary observation network. Therefore, synthetic data were generated at the highest sampling elevation of the nine towers that were collecting continuous high-precision calibrated CO_2 measurements in North America in June of 2004 (Figure 5.1, Table 5.2). A full set of synthetic measurements without data gaps from June 1 to July 8 were generated by multiplying 3-hourly CO_2 surface flux estimates ($\mathbf{s}$) from a biospheric model by the atmospheric transport matrices ($\mathbf{H}$).

Table 5.2: Measurement locations used in the inversions.

Tower	*Location*	*Coordinates*	*Height*	*Maintained by*	*Type*
LEF	Park Falls, Wisconsin	45.93N, 90.27W	396 m	NOAA/GMD	Tall
WKT	Moody, Texas	31.32N, 97.33W	457 m	NOAA/GMD	Tall
BRW	Barrow, Alaska	71.32N, 156.60W	10 m	NOAA/GMD	MBL
SBL	Sable Island, Nova Scotia	43.93N, 60.02W	25 m	Met Service Canada	MBL
AMT	Argyle, Maine	45.03N, 68.68W	107 m	NOAA/GMD	Short
ARM	Norman, Oklahoma	36.62N, 97.50W	60 m	U.S. Dept. of Energy	Short
CDL	Candle Lake, Saskatchewan	53.99N, 105.12W	30 m	Met Service Canada	Short
FRD	Fraserdale, Ontario	49.84N, 81.52W	40 m	Met Service Canada	Short
HFM	Petersham, Massachusetts	42.54N, 72.17W	30 m	Harvard University	Short

The biospheric fluxes used in this study are taken 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). CASA-GFEDv2 was chosen because it is a well-accepted model that has been used for specifying prior flux estimates in several synthesis Bayesian inversion studies (e.g. Baker et al., 2006, Peters et al., 2007), although the choice of biospheric model is flexible here, given that the aim of a synthetic data inversion is to assess the accuracy of the setup relative to a given set of prescribed fluxes. The monthly-average CASA-GFEDv2 Net Ecosystem Exchange (NEE) values were temporally downscaled to a 3-hourly resolution in order to test the ability of the inversion setup to accurately recover diurnally-varying fluxes. This was accomplished using the method of Olsen & Randerson (2004), which is based on net shortwave radiation and near-surface temperature data from the NASA Global Land Data Assimilation System (GLDAS; Rodell et al., 2004). These downscaled 3-hourly CASA-GFEDv2 fluxes, shown in Figure 5.3 at the aggregated monthly scale, represent the "truth" to which inversion results are compared.

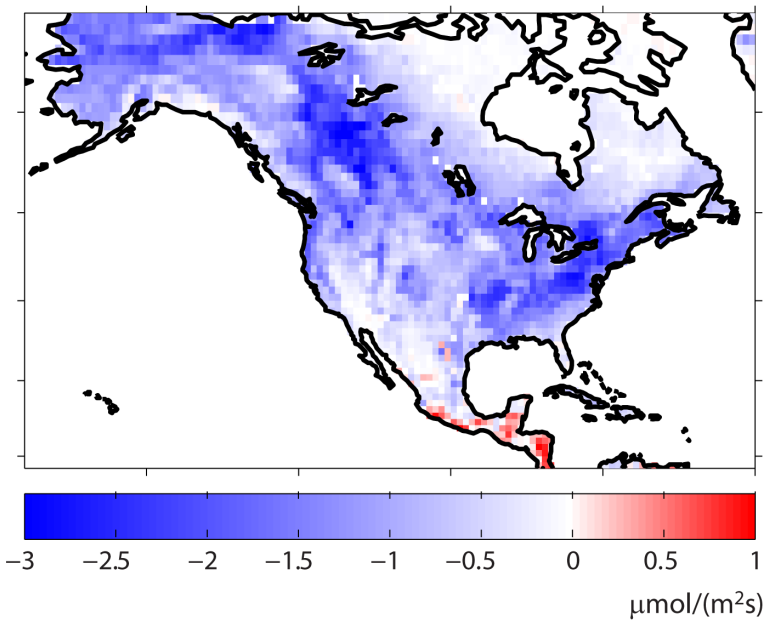

Figure 5.3: "True" CASA-GFEDv2 fluxes, aggregated to the monthly scale.

The vector of modeled observations (i.e. $\mathbf{Hs}$), obtained by multiplying the "true" 3-hourly fluxes by the concentration footprints, is first generated at the hourly resolution corresponding to the STILT particle releases for the nine tower locations. Then, the synthetic observation vectors are averaged to a 3-hourly timescale. Sensitivity tests were additionally performed for inversions using daily and 8-day average concentration vectors. These tests revealed that using higher temporal resolution observations yielded superior flux estimates, consistent with Law et al. (2002), who found that using 4-hourly measurements relative to more coarsely-averaged observations helped to reduce biases over the Australasian subcontinent when fluxes were estimated at a sufficiently fine spatial resolution. As a result, only cases considering 3-hourly averaged observations are presented here.

Finally, for inversions that simulate the effect of transport model error, three sets of inversions using different realizations of uncorrelated errors added to the synthetic measurements were conducted. The variance of these errors remains the same across realizations, and corresponds to the expected magnitude of model-data mismatch seen in real measurements for each tower, as discussed in Section 5.3.1.

5.2.4 Use of night-time measurements

Due to stable conditions, night-time measurements taken from shorter towers within the nocturnal boundary layer provide little information about fluxes over large aggregated spatial and temporal scales (Haszpra, 1999). In addition, meteorological fields used in transport models have difficulty reliably simulating the height of the night-time planetary boundary layer (PBL), or the sharp gradient across it, which can lead to biased flux estimates from inversions using night-time measurements from within the PBL (Geels et al., 2007). For example, Gerbig et al. (2008) found biases of up to 50% in night-time PBL height in a study using high-resolution winds from the European Centre for Medium-Range Weather Forecasts (i.e. ECMWF, available from http://data.ecmwf.int/data/); comparison of the high resolution WRF wind fields (T.

Nehrkorn, personal communication) used in this study with wind profiler PBL-height measurements yielded a similar conclusion.

Given their local footprints, the use of night-time near-surface measurement data in regional inversions could lead to higher aggregation errors near the towers, relative to those caused by afternoon measurements sampling well-mixed air. In addition, biases in night-time PBL height would affect future real-data inversions. For these reasons, only afternoon measurements are included here for the five "Short" towers (see Table 5.2) that are consistently within the nocturnal boundary layer. For these towers, which are all in the Eastern or Central Standard Time zones, "afternoon" was considered to be 1800-2400 UTC. In contrast, all 24 hours of atmospheric data were included in the inversions for the four tall (≥ 400m) or marine boundary layer towers ("Tall" or "MBL" in Table 5.2). At these towers, observations sample relatively well-mixed air throughout the diurnal cycle, and therefore night-time measurements are assumed to be better-represented by the WRF/ STILT model relative to the Short towers. This was qualitatively confirmed by comparing 24 hours of actual observations at the tallest sampling levels of these towers to those from transported CASA-GFEDv2 fluxes.

Two sensitivity tests were performed to evaluate the choice of including night-time data for Tall and MBL towers and excluding them for Short towers. First, night-time data were included for the Short towers, such that 24 hours of measurements were used for all sampling locations. Second, night-time data were excluded for the Tall/MBL towers, such that only afternoon measurements were used for all nine towers. (For the MBL towers, afternoon values were shifted to reflect local time zones.) Overall, such experiments help to assess biases associated with the use of night-time measurements, relative to the potential additional constraint on fluxes they can provide.

5.2.5 Model of the trend ($\mathbf{X\beta}$)

A very simple model of the trend is applied in the current study for all inversions, analogous to those used in Michalak et al. (2004) and Mueller et al. (2008), where no additional auxiliary environmental variables are included in the model of the trend. As discussed in Chapter 3, the flux estimates ($\hat{\boldsymbol{s}}$) are a composite of the inferred trend ($\mathbf{X}\widehat{\boldsymbol{\beta}}$) and a spatiotemporally-correlated stochastic component, such that any variability in the atmospheric signal not captured by the trend can be still be recovered through the stochastic component of the best estimate. For the F4d inversions, $\mathbf{X}$ is represented as a vector of ones, where the corresponding drift coefficient ($\widehat{\boldsymbol{\beta}}$) represents the mean value in space and time of fluxes across all grid-cells. For the two temporal resolutions resolving the diurnal cycle (F4d-diurnal and F3hr), the $\mathbf{X}$ matrix is instead structured to allow for eight spatial means defined for each 3-hourly bin of the diurnal cycle. Longitudinal gradients that could capture the changing day/ night boundary across the continent for different UTC time intervals were also considered for these inversions, but ultimately discarded because they did not help to improve flux estimates.

5.2.6 Covariance matrices ($\boldsymbol{Q}$ and $\boldsymbol{R}$)

Model-data mismatch errors are assumed uncorrelated in space and time, yielding a diagonal matrix $\mathbf{R}$, as is typical in most inversion studies. A different variance was used for each measurement tower, given results from initial tests showing significantly reduced biases in inferred fluxes as compared to an inversion using only two separate variances for Tall/MBL and Short towers.

In contrast to the diagonal structure of $\mathbf{R}$, the covariance matrix $\mathbf{Q}$ contains off-diagonal entries describing the spatial and/or temporal correlation of the flux deviations from the model of the trend $\mathbf{X\beta}$. Because of the simple models of the trend used in this study (Section 5.2.6), the deviations represent residuals from a constant mean, and therefore $\mathbf{Q}$ describes the covariance of the fluxes themselves.

The estimated fluxes for this study are sorted first in space, and then in time. Therefore, if only spatial covariance were considered, $\mathbf{Q}$ would be a block diagonal

matrix, with each block describing the correlation between grid-scale fluxes for each time period of the inversion. When temporal covariance is additionally considered, the off-diagonal blocks in $\mathbf{Q}$ contain diagonal entries describing the correlation among grid-cells with themselves over time. Finally, if cross spatial-temporal covariance is included, the off-diagonal blocks in $\mathbf{Q}$ become full, and they describe the spatial covariance between fluxes across different time periods.

In the current study, preliminary tests showed that including full spatiotemporal covariance between grid-scale fluxes helped to recover accurate uncertainty bounds for recovered fluxes, especially at spatially and temporally aggregated scales. Therefore, cross spatial-temporal covariance was included in $\mathbf{Q}$ for all flux temporal resolutions. However, for the two flux resolutions resolving the diurnal cycle, spatial-temporal covariance is only assumed for the same 3-hourly interval across days or periods, but not within the diurnal cycle. For example, grid-scale fluxes from 0-300 UTC are correlated with fluxes from 0-300 UTC in neighboring days or periods, but never with fluxes from 300-600 UTC.

The correlation structure in $\mathbf{Q}$ is modeled using a covariance function that varies in space and time as a function of separation distance. Here, as in Michalak et al. (2004), we use an isotropic exponential decay model:

$$\boldsymbol{Q}\left(h_x, h_t \middle| \sigma_{\boldsymbol{Q}}^2, l, \tau\right) = \sigma_{\boldsymbol{Q}}^2 exp\left(-\frac{h_x}{l}\right) \exp\left(-\frac{h_t}{\tau}\right) \qquad (5.1)$$

where h_x and h_t are the separation distances between grid cells in space and time, respectively, l is the spatial correlation range parameter, τ is the temporal correlation range parameter, and $\sigma_{\mathbf{Q}}^2$ is the asymptotic variance of fluxes at large separation distances. The correlation length for an exponential model is approximately $3l$ or 3τ.

Multiple variance parameters $\sigma_{\mathbf{Q}}^2$ were initially considered for different times of the day for those inversions resolving the diurnal cycle (i.e. F4d-diurnal and F3hr). For example, the underlying "true" CASA-GFED fluxes are significantly more variable during the day-time compared to the night. However, the use of multiple variance parameters

resulted in only small changes to the inferred fluxes, and in some areas biased the results. Therefore, for simplicity, only one flux variance parameter was used for all inversions, regardless of flux temporal resolution.

Covariance parameters were estimated using the Restricted Maximum Likelihood (RML) approach, described in more detail in Chapter 3 (Section 3.2). More specifics on the implementation of RML in this study are also included below in Section 5.3.1.

5.3 Covariance parameter optimization

This section describes the approach taken to estimate covariance parameters in this study, as well as an analysis of the inferred parameters.

5.3.1 Setup for testing RML optimization with atmospheric data

To infer unbiased fluxes with accurate uncertainty estimates, it is important to correctly specify the flux covariance parameters (Gerbig et al., 2006), as well as the model-data mismatch variances. The RML approach provides a way to statistically optimize these parameters using the atmospheric data in an inverse setup (henceforth referred to as RML-Inv, see equation 3.12). If the recovered covariance parameters can be shown to yield accurate flux estimates, then this approach eliminates the need to use proxy methods for estimating covariance parameters. While the RML approach for covariance parameter optimization was previously demonstrated in Michalak et al. (2004) to perform well with synthetic data experiments for the global scale, it is investigated here for regional inversions using continuous data.

To test performance, parameters inferred using RML-Inv for each flux temporal resolution are compared to reference values, derived using the underlying true fluxes, as described below. In addition, the impact of using RML-Inv parameters on estimated fluxes is investigated in Section 5.4.

In order to derive the reference values, the covariance parameters for $\mathbf{Q}$ are estimated by implementing RML directly on the "true" underlying CASA-GFEDv2 fluxes

($\boldsymbol{s}$, see equation 3.11). This approach will henceforth be referred to as RML-Krig, where "Krig" refers to the kriging setup of RML. Because RML-Krig does not use the atmospheric measurements, the RML-Krig parameters are not affected by the simulation of transport error.

In order to derive "true" model-data mismatch variances in $\mathbf{R}$, an approach other than RML-Krig must be applied using the true fluxes transported forward to the measurement locations. In the current study, fluxes are estimated at the native spatial resolution of the true fluxes (i.e. 1°x1°). Therefore, there is technically no spatial aggregation error, and without adding random noise to the measurements to simulate transport error, the model-data mismatch variances in $\mathbf{R}$ are exclusively determined by temporal aggregation error. These errors can be directly calculated as the variance of the difference between two synthetic data vectors: the observations used in the inversion (generated using 3-hourly fluxes), and a second set of observations generated using fluxes pre-averaged to coarser timescales (i.e. the 4-day or 4-day diurnal cycle) and then multiplied by aggregated transport matrices ($\mathbf{H}$). Temporal aggregation error is technically zero in this study when estimating 3-hourly fluxes, although a floor of 0.01 ppm^2 is set for the model-data mismatch variance in all inversions using this flux resolution.

For inversions that consider transport model errors, the variance of the noise added to the measurements is added to the temporal aggregation error variance to arrive at the "true" total model-data mismatch variance. The variances of the added noise were determined per tower in the following manner. RML-Inv was first used to estimate $\mathbf{Q}$ and $\mathbf{R}$ parameters with actual atmospheric measurements for June 2004. Then, the difference between these model-data mismatch variances and the RML-Inv inferred temporal aggregation errors (with synthetic data) were taken as a measure of the magnitude of residual model-data mismatch for each tower, most of which is likely attributable to transport model error. This procedure was repeated for each of the examined flux resolutions, and then these inferred differences by tower were used as the variance of the random noise added to the synthetic measurements for the

transport error analyses.

5.3.2 Comparison of reference covariance parameters by flux temporal resolution

The RML-Krig $\mathbf{Q}$ parameters and the "true" $\mathbf{R}$ covariance parameters are shown in Tables 5.3 and 5.4 for each of the examined flux resolutions. Because these parameters were inferred using the true underlying fluxes, they are henceforth referred to as reference values, for later comparison with the RML-Inv parameters inferred with the synthetic measurements. However, first we compare the reference parameters themselves across flux temporal resolutions.

Table 5.3: Inferred parameters using RML-Krig and RML-Inv for the a priori flux covariance matrix ($\mathbf{Q}$). RML-Inv parameters were estimated using observations with and without simulated transport error. The cell shading indicates the factor by which RML-Inv parameters differ from the RML-Krig, a.k.a. reference, values:

(no fill) ≤ 2× | 2× to 3× | 3× to 4× | 4×

Q	**F4d**			**F4d-diurnal**			**F3hr**		
	RML-Krig	RML-Inv		RML-Krig	RML-Inv		RML-Krig	RML-Inv	
		Perfect trans.	Trans. error		Perfect trans.	Trans. error		Perfect trans.	Trans. error
σ^2_q $(\mu mol/(m^2 s))^2$	1.0	97.1	25.8	11.1	35.5	27.3	13.4	24.3	20.7
l (km)	610	0	57	809	389	363	601	661	528
τ (days)	6.8	2.6	6.9	80.7	9.2	9.6	8.6	2.7	3.5

The RML-Krig $\mathbf{Q}$ parameters show that the spatial correlation ranges (l) are not substantially different across temporal resolutions, but that the overall variance ($\sigma^2_{\mathbf{Q}}$) of the fluxes decreases as the estimated temporal resolution becomes coarser. This is expected, as more of the short-term variability in the spatial flux distribution is averaged out. The temporal correlation range (τ) is much longer for the 4-day diurnal cycle relative to the other timescales, although this is most likely an unreliable value given

that the calculated correlation length is much longer than the one-month time period of analysis.

The "true" model-data mismatch variances in $\mathbf{R}$ for inversions without transport error represent temporal aggregation errors, and they become higher, as expected, as the flux temporal resolution becomes coarser. This is the essence of aggregation error, where averaging out the "true" temporal variability in the fluxes and then transporting them forward to the sampling locations cannot properly reproduce the measured concentrations. For a given temporal resolution, these aggregation errors also tend to be higher for towers in highly active biospheric regions (e.g. LEF, AMT and HFM), where temporal variability in nearby fluxes has a strong influence on measured concentrations. The "true" model-data mismatch with simulated transport error is increased by the magnitude of the random noise added to the measurements, which varies from a standard deviation of about 0.5 ppm for CDL to about 5 ppm for AMT.

5.3.3 Results of RML-Inv Optimization

Estimated RML-Inv parameters for $\mathbf{Q}$ and $\mathbf{R}$ are also shown in Tables 5.3 and 5.4. The RML-Inv values for inversions with transport error represent the average parameters inferred using three realizations of random noise. Although RML-Inv parameters are compared here with the RML-Krig values (for $\mathbf{Q}$) and "true" model-data mismatch variances (for $\mathbf{R}$) to assess the relative ability of the atmospheric data to recover covariance parameters, the ultimate concern is the impact of these parameters on the inversions, results of which are presented in Section 5.4.

*Table 5.4: "True" and RML-Inv inferred variances by tower for the model-data mismatch matrix (**R**). Both "True" and RML-Inv results are shown as calculated using observations with and without simulated transport error. The cell shading is the same as in Table 5.3, indicating the factor by which RML-Inv parameters differ from the "true", or reference, values.*

R	**F4d**				**F4d-diurnal**				**F3hr**			
(ppm^2)	Perfect transport		Transport error		Perfect transport		Transport error		Perfect transport		Transport error	
	"True"	RML-Inv	"True"	RML-Inv	"True"	RML-Inv	"True"	RML-Inv	"True"	RML-Inv	"True"	RML-Inv
LEF	10.7	2.7	13.8	7.9	2.2	0.1	4.7	2.7	0	0	1.8	1.9
WKT	5	2.3	18	15.7	0.7	0.1	10.5	7.9	0	0	10.1	8.8
SBL	6.2	0.1	10.6	5.6	2.7	0.1	7.3	5	0	0	4.6	4.3
BRW	0.2	0	1.3	1.1	0.1	0	1.2	1.1	0	0	1.1	1.2
ARM	3.5	0.1	14.9	13.1	1.2	0.4	12.4	10.3	0	0	11.1	11.2
HFM	15.6	5.2	34.4	25.7	8.8	6.8	24.4	17.7	0	2.2	23.5	23.2
AMT	9.5	3.9	23.4	17.6	4.8	1.3	21.5	14.9	0	0.3	11.6	13.4
FRD	3.4	0	11.7	7.4	2.4	0	11.1	7.1	0	0	8.8	8.6
CDL	2.9	0	3.6	1	1.2	0	1.4	0	0	0	0	0

The flux variance parameter ($\sigma_{\mathbf{Q}}^2$) is higher than the reference value for all temporal resolutions. This is most likely due to the fact that the majority of towers are sited in biospherically active regions, which have above-average flux variability as compared to the continent as a whole. Surprisingly, the addition of simulated random transport error helps to bring the flux variance closer to the reference value, but this may be due to the transport errors obscuring some of the "true" flux variability that would otherwise be seen through the measurement data.

Overall, estimating covariance parameters with the atmospheric data and the coarse flux resolution (F4d) appears to yield consistently unreliable parameter estimates, whereas the RML-Inv approach with the other two temporal resolutions yields results that are more consistent with the reference values. The impact of these inferred parameter estimates on inversion results is explored in the next section.

5.4 Inversion Experiments

This section describes the setup and results of inversions used to test the impact of covariance parameter optimization methods, flux temporal resolutions, and other inversion assumptions and data choices.

5.4.1 Inversion setups and diagnostics

After the covariance parameter optimization analysis, a series of inversions was run to test the impact on inferred fluxes and uncertainties of a) varying the temporal flux resolution (shown in Table 5.1), (b) using inferred RML-Inv vs. the reference covariance parameters, and c) including simulated random transport error. All combinations of a), b) and c) yielded 24 inversions, i.e. 2 sets of covariance parameters for each of 3 flux temporal resolutions, once with perfect transport and then with three different realizations of random transport error. In addition, sensitivity test inversions with perfect transport were run to test the impact of excluding cross spatial-temporal covariance in $\mathbf{Q}$, and the inclusion or exclusion of night-time data.

The *a posteriori* flux estimates from all inversions, as well as the true CASA-

GFEDv2 fluxes, were averaged to a monthly timescale in order to compare results at an aggregated scale relevant for carbon-cycle science. The inferred grid-scale fluxes are compared using two quantitative metrics. First, the root mean square error (RMSE) (e.g. Law et al. (2002)) between the true and estimated fluxes was calculated at the native 1°×1° spatial resolution for all land grid-cells across the continent (Figure 5.4). Second, the accuracy of the estimated *a posteriori* uncertainties (from equation 3.5) was evaluated by calculating the percent of 1°×1° true fluxes that fall within two standard deviations of the estimated fluxes (Table 5.5). Ideally, 95% of fluxes should fall within this interval. Values significantly below 95% indicate an underestimation of the true *a posteriori* uncertainties. The results of this second metric are compared for inversions with and without cross spatial-temporal covariance in $\mathbf{Q}$, in order to examine the impact of accounting for temporal covariance on the recovered flux uncertainties.

Inversion results are also compared qualitatively by examining the spatial patterns of inferred fluxes to those of the true fluxes (Figures 5.5, 5.6, and 5.7). Finally, monthly fluxes and uncertainties are aggregated to seven ecoregions (Figure 5.8) as well as to the North American continent (Figure 5.9). These ecoregions are loosely defined based on the work of Olson (2001), and represent large, mostly contiguous, regions with similar climate, land cover and land use. An area-weighted RMSE at the ecoregion scale was also calculated.

5.4.2 Results of grid-scale diagnostics

Figure 5.4 shows the grid-scale RMSE's for the six perfect transport inversions (using two sets of covariance parameters and grouped by the three examined flux resolutions). With the reference covariance parameters, there is little difference in continental grid-scale RMSE among the three flux resolutions, although the 3-hourly resolution shows a slight advantage. Using the RML-Inv parameters, inversion performance degrades for coarser estimation timescales, consistent with the fact that the recovered RML-Inv covariance parameters became farther from the reference values as fluxes were temporally aggregated (Section 5.3.3). If the inversion were to

infer the exact mean monthly flux across the continent with no spatiotemporal variability, the RMSE would be 0.72 µmol/(m^2s). Therefore, all inversions, except for the F4d case with RML-Inv parameters, perform better than this baseline value.

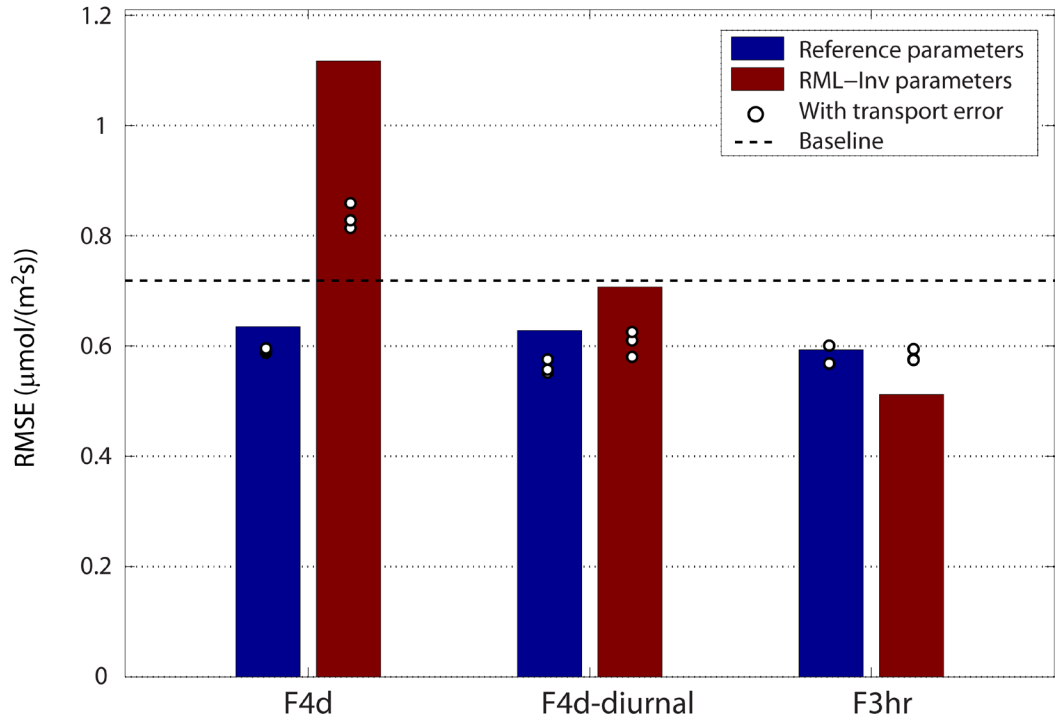

Figure 5.4: Root Mean Square Error (RMSE) between estimated and "true" grid-scale fluxes, aggregated to monthly averages, for inversions with two sets of covariance parameters and three flux temporal resolutions. Bars show the RMSE for inversions without transport error, and the white dots show the RMSE for inversions conducted with three different realizations of simulated transport error. The baseline RMSE, as described in Section 4.2, represents the value associated with inferring a perfect flat mean monthly flux across the continent.

The RMSE corresponding to the three realizations of simulated transport error for each inversion setup are also shown in Figure 5.4. The grid-scale fluxes (results not shown) from these inversions show that transport error has the effect of damping down the variability in the inferred flux signal. For the two timescales with temporal aggregation error (i.e. F4d and F4d-diurnal), this has a positive impact on the RMSE, with the difference being more pronounced with the RML-Inv parameters due to the

improvement in quality of the covariance optimization for these cases. In contrast, for the F3hr case, simulated transport error has a minimal impact on the grid-scale RMSE. While these results are promising in that random transport errors are unlikely to bias inferred grid-scale fluxes, it is not clear whether this result would hold true with more realistic systematic, non-random transport errors.

The fraction of true fluxes lying within two standard deviations (equation 3.5) of the estimated fluxes is presented at the monthly timescale for the inversions with no transport error in Table 5.5, for both covariance parameter optimization methods, and with and without temporal covariance included in $\mathbf{Q}$. The inclusion of temporal covariance is found to be important for obtaining accurate *a posteriori* uncertainties when using the reference covariance parameters for all flux resolutions. The same result holds using the RML-Inv parameters with the 3-hourly flux resolution, although for the F4d-diurnal and F4d cases, accurate grid-scale uncertainties can be obtained with spatial-only covariance due to the overestimation of the flux variance parameter in $\mathbf{Q}$. Results with simulated transport error are not shown in Table 5.5, although the same conclusions hold. Overall, these results highlight that accounting for both spatial and temporal flux covariance yields accurate *a posteriori* uncertainty bounds at the grid-scale much more reliably than accounting for only spatial correlations.

Table 5.5: Percent of true fluxes falling within two standard deviations of the a posteriori grid-scale monthly flux estimates. Results are shown without added transport errors for the three flux temporal resolutions, two sets of covariance parameters, and with and without assumed temporal flux covariance in $\mathbf{Q}$.

	Reference parameters		**RML-Inv parameters**	
	Spatial covariance only	Spatiotemporal covariance	Spatial covariance only	Spatiotemporal covariance
F4d	66%	88%	100%	100%
F4d-diurnal	80%	96%	98%	99%
F3hr	49%	93%	63%	95%

5.4.3 Inferred grid-scale spatial patterns

Grid-scale maps of monthly-averaged flux estimates are shown in Figures 5.5, 5.6 and 5.7 for inversions performed with the three flux resolutions and two sets of covariance parameters, without any simulated random transport error. For comparison, the true monthly-averaged fluxes are shown in Figure 5.3. Overall, the inversions detect large-scale patterns of sources and sinks consistent with the true underlying fluxes, with significant sinks recovered from the eastern United States to northwest Canada and Alaska. As expected, the inferred fluxes show significantly less overall variability relative to the true fluxes, due to the sparse atmospheric network and the absence of auxiliary environmental variables within the trend (e.g. Gourdji et al. 2008).

The F4d inversions (Figure 5.5) perform least well in capturing the true grid-scale spatial patterns. Fluxes estimated using the reference covariance parameters remain close to their mean value for the continent with little spatial variability. In addition, high normalized errors exist near the WKT, ARM and SBL measurement towers, most likely due to temporal aggregation errors near the sampling locations. High errors are also seen in northwest Canada, where the strong sinks in this region fall outside of the areas well-constrained by the atmospheric measurements (Figure 5.2). With the RML-Inv parameters, the opposite problem occurs such that there is unrealistic spatial variability in the recovered fluxes associated with the artificially high flux variance parameter in $\mathbf{Q}$. These results are consistent with those presented in Section 5.3.3, confirming that estimating fluxes directly at highly aggregated temporal scales is not an optimal setup for regional inversions.

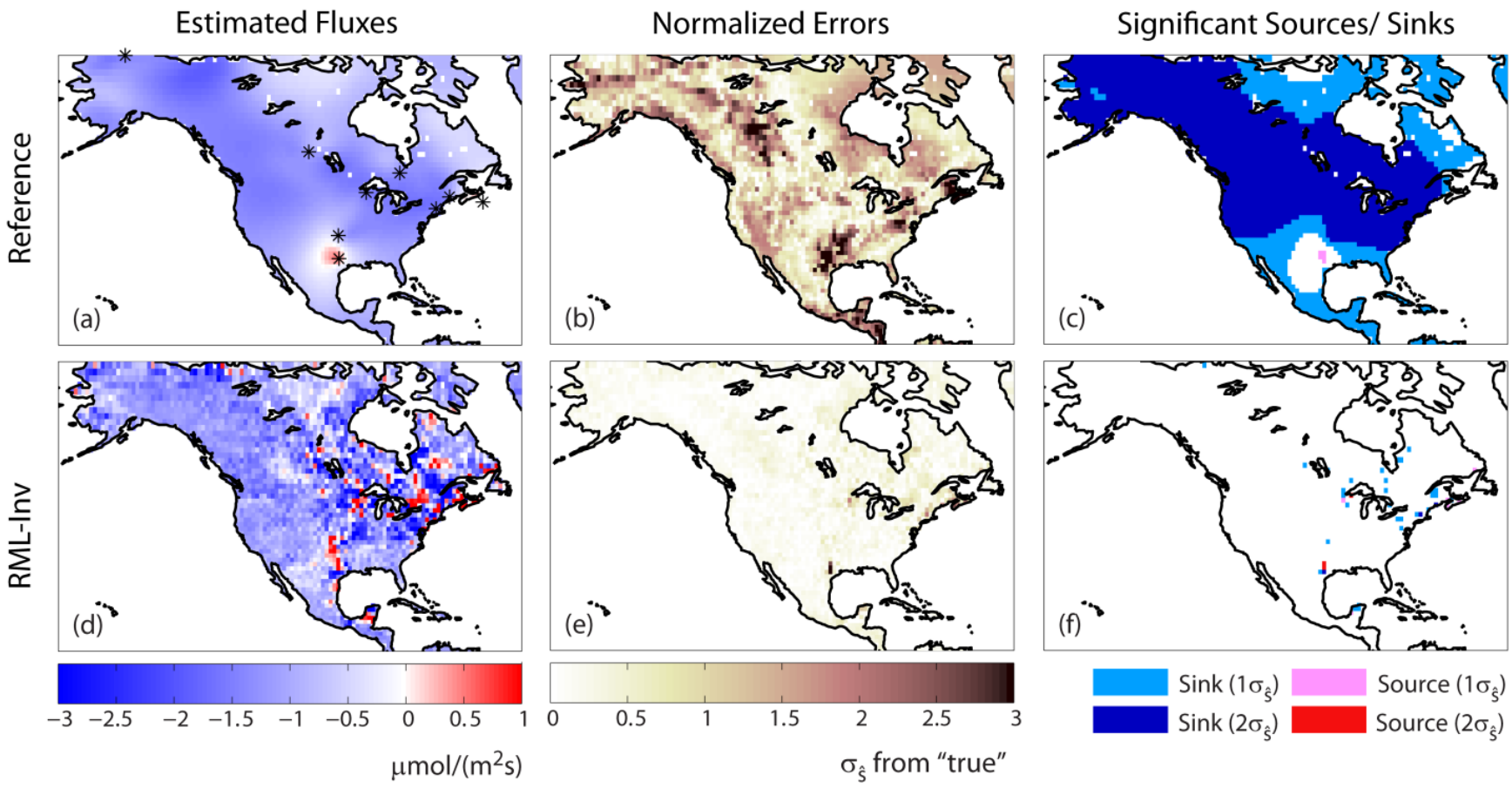

Figure 5.5: Monthly grid-scale fluxes estimated from the inversion using the 4-day average (F4d) temporal resolution. The first row presents results obtained using the reference covariance parameters, and the second row shows results obtained using RML-Inv parameters. The sub-figures show a posteriori fluxes aggregated to monthly averages (a, d), errors relative to the true fluxes (Figure 5.3) normalized by their a posteriori standard deviations (b, e), and grid-scale significant sources and sinks at $1\sigma_{\hat{s}}$ and $2\sigma_{\hat{s}}$ (c, f).

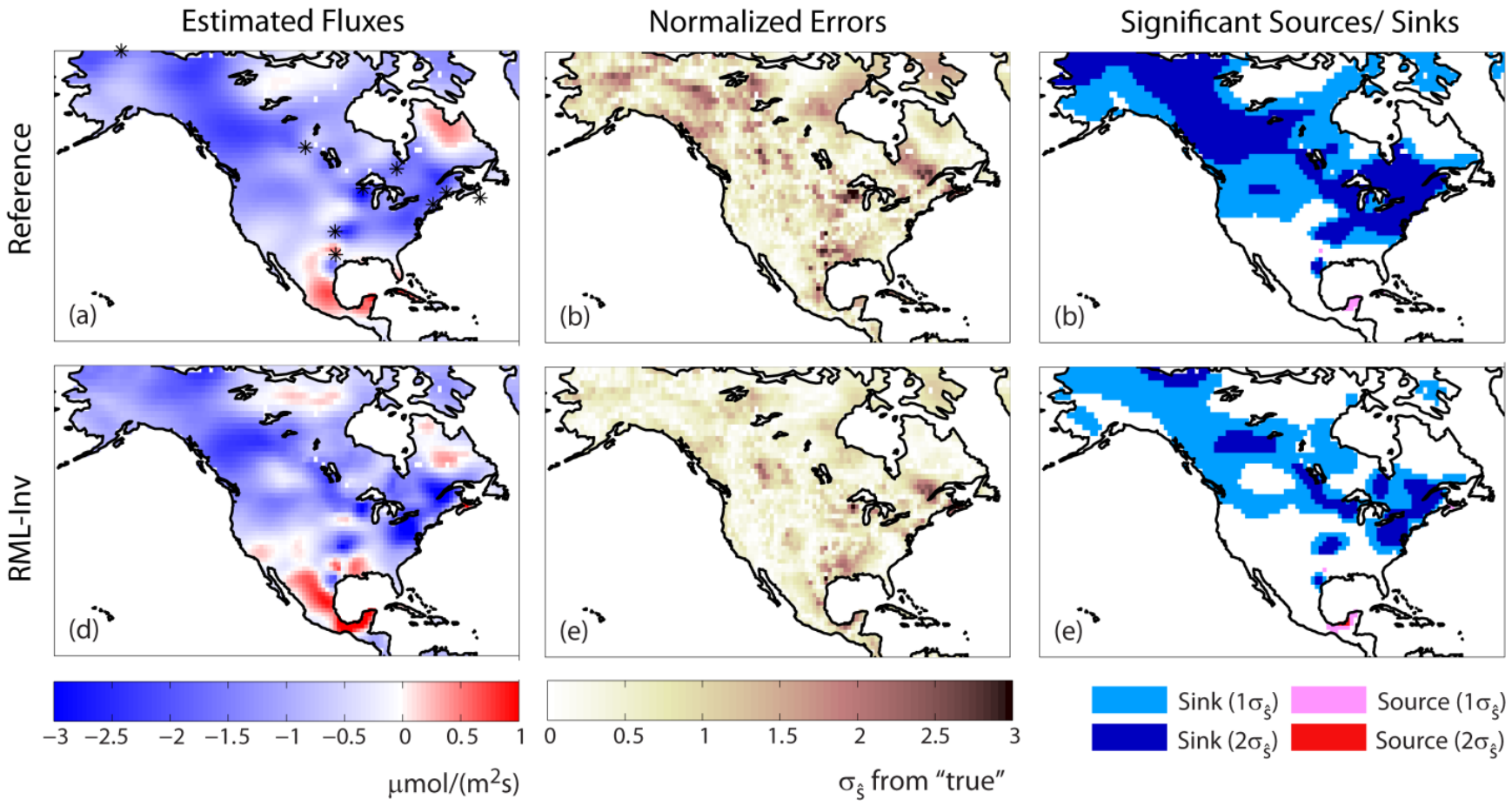

Figure 5.6: The same as Figure 5.5, except using the 4-day average diurnal cycle (F4d-diurnal) temporal flux resolution.

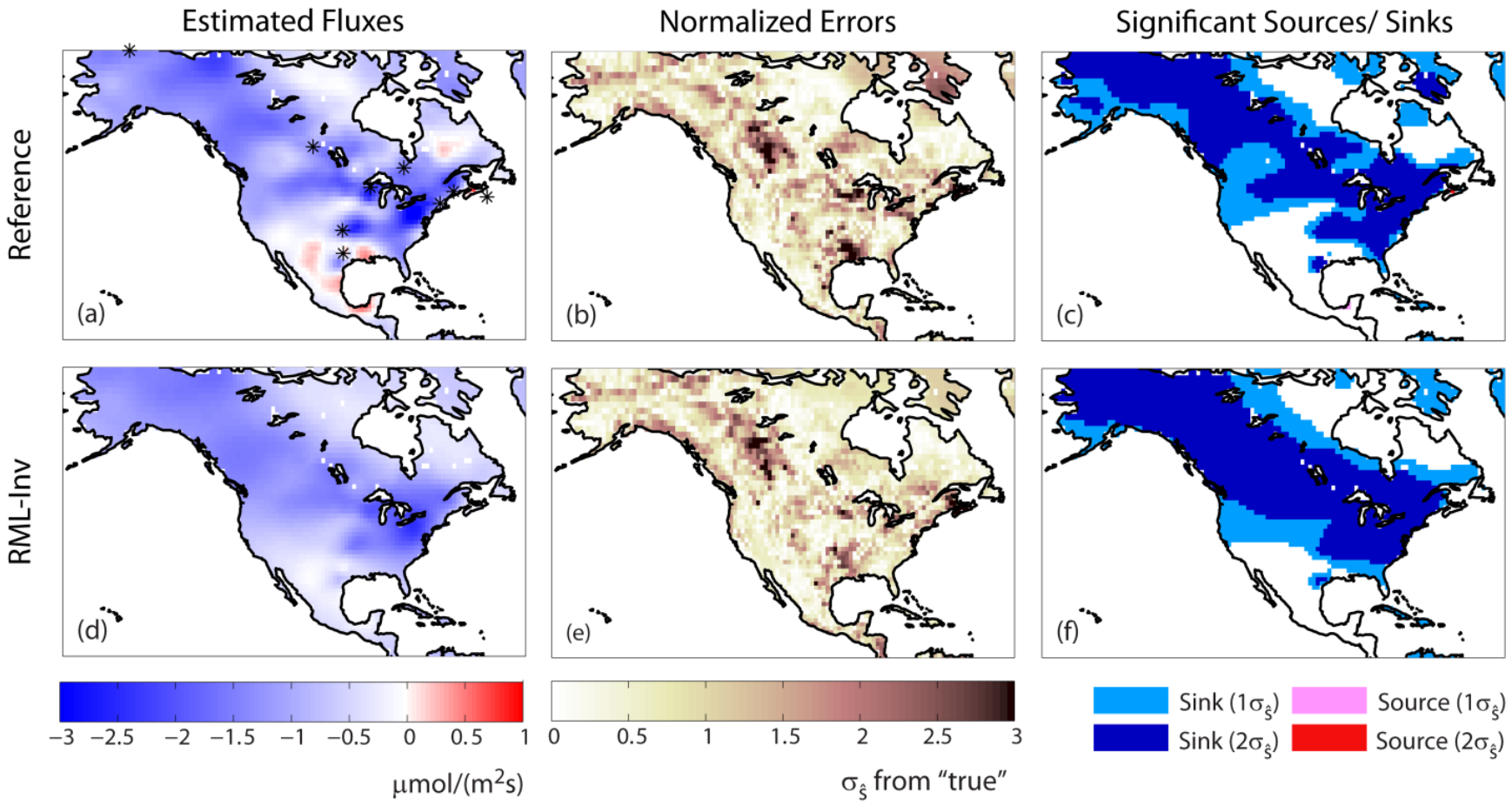

Figure 5.7: The same as Figure 5.5, except using the 3-hourly (F3hr) temporal flux resolution.

The F4d-diurnal inversions (Figure 5.6) show more realistic spatial variability in the fluxes. Sources are now properly recovered in Central America, while normalized errors are reduced in all areas. Also, fluxes recovered using both sets of covariance parameters yield consistent results with reasonable grid-scale spatial patterns. This indicates that, although the covariance parameters for the F4d-diurnal case inferred using the atmospheric data (RML-Inv) differed from the reference covariance parameters in some cases (Tables 5.3 and 5.4), the RML-Inv parameters can still be used to recover fluxes of comparable quality to those obtained using idealized covariance parameters. The *a posteriori* uncertainties are more affected by the use of the RML-Inv parameters than the fluxes, as reflected in the lower normalized errors and fewer significant sources and sinks relative to the inversion using the reference parameters.

The F3hr inversions (Figure 5.7) also yield realistic spatial variability, especially when using the reference parameters. With the RML-Inv parameters, the spatial variability is slightly reduced, although as with the F4d-diurnal case, inferred fluxes are similar using the two sets of parameters. Also, given the more realistic recovered uncertainties with this resolution as compared to the other cases (Table 5.5), the largest number of significant sources and sinks are recovered at the grid-scale with the RML-Inv parameters for this case.

Overall, this comparison of the grid-scale flux maps demonstrates that an inversion that resolves the diurnal cycle of the fluxes (i.e. F4d-diurnal and F3hr) can recover reasonably accurate grid-scale spatial patterns across the continent using only a 9-tower measurement network and no auxiliary process-based information. Also, as previously shown in Figure 5.4, the quality of inferred fluxes is preserved even when covariance parameters are estimated from the available atmospheric data, and not assumed known *a priori*. Inversions that estimate fluxes at coarser timescales (i.e. F4d) that average out the diurnal cycle do not perform nearly as well, consistent with the covariance parameter conclusions presented in Section 5.3.3.

5.4.4 Results at aggregated ecoregion scale

Figure 5.9 presents estimated fluxes and their uncertainties, from the inversions using covariance parameters inferred with the atmospheric observations (RML-Inv), aggregated *a posteriori* to the monthly average, ecoregion (Figure 5.8) and continental scales. The RML-Inv parameters were used, because this is most consistent with what would be possible in future real-data inversions. Also, results using the reference covariance parameters are very similar at this aggregated scale to those presented in Figure 5.9.

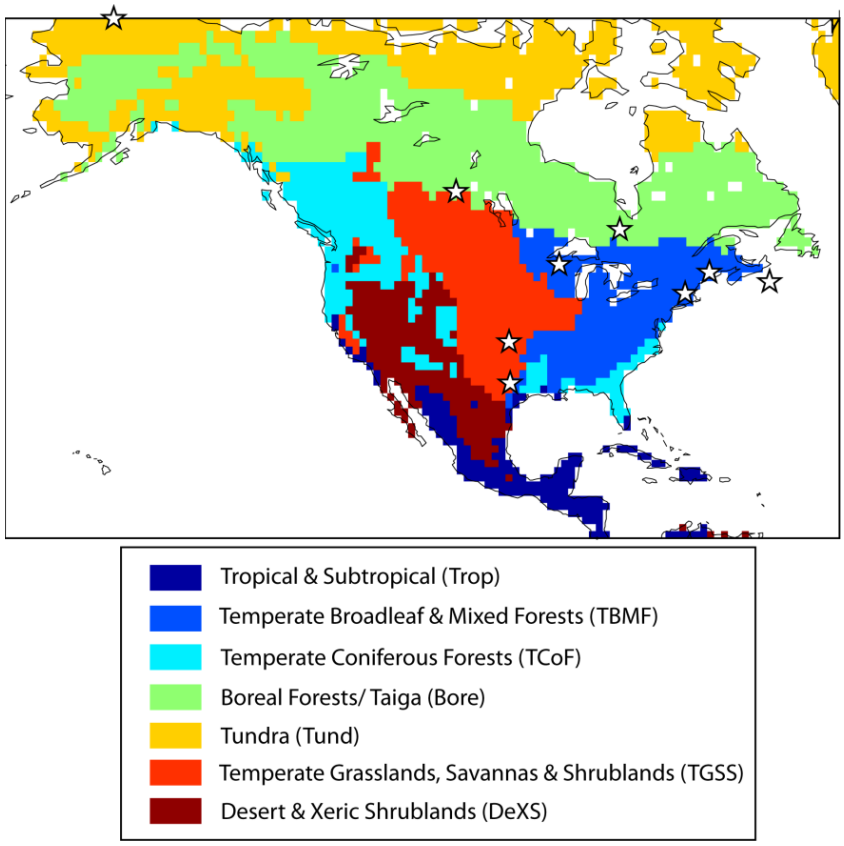

Figure 5.8: Seven ecoregions (modified from Olson (2001)) used for analyzing inversion results at spatially aggregated scales. Stars represent nine measurement locations.

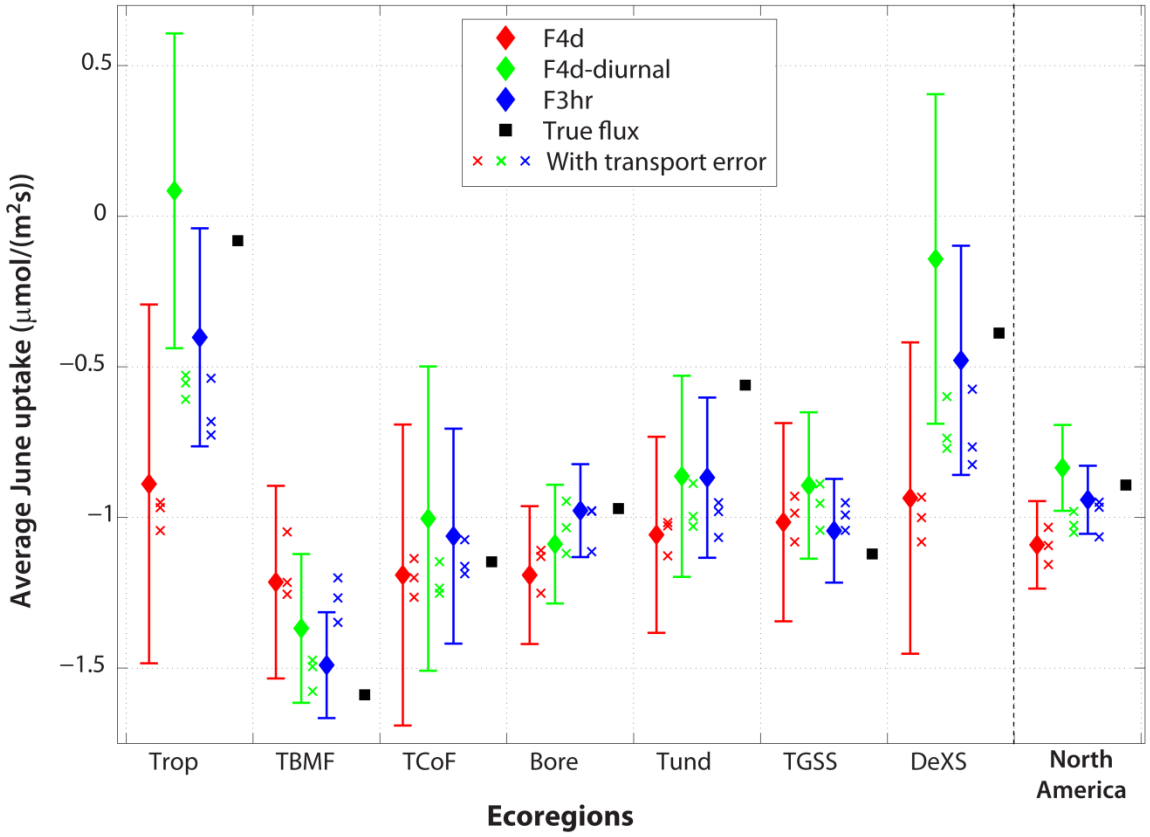

Figure 5.9: Estimated fluxes from inversions with no transport error using RML-Inv covariance parameters, aggregated to monthly average ecoregion (Figure 5.8) and continental scales. Associated error bars represent 95% uncertainty bounds. Estimated fluxes from inversions using three different realizations of transport error are also shown for each temporal resolution. Uncertainty bounds are not shown for these inversions for simplicity, but were similar in magnitude to their equivalents with no transport error.

At the ecoregion scale, the inversions resolving the diurnal cycle (F4d-diurnal and F3hr) are seen to yield more accurate fluxes, which is also confirmed by the RMSE's at this aggregated scale (0.03, 0.04 and 0.16 μmol/(m^2s) for the F3hr, F4d-diurnal, and F4d cases, respectively). The relative performance of inversions using these three temporal resolutions is particularly evident in the better-constrained ecoregions, such as the Temperate Broad & Mixed-leaf Forest and Boreal Forest, where temporal aggregation error has the most impact on fluxes due to their a) proximity to the towers and b) strong flux variability. In addition, the F3hr and F4d-diurnal inversions also infer more realistic ecoregion fluxes in the far-field, e.g. in the Tropics and the Desert & Xeric Shrubland. In

contrast, the F4d case yields aggregated fluxes that remain close to the mean continental flux across all ecoregions. This last result is interesting in that high grid-scale spatial variability for this case (seen in Figure 5.6) using the RML-Inv parameters did not translate into large differences in spatially aggregated ecoregion-scale fluxes.

For the F3hr and F4d-diurnal cases, the 95% uncertainty bounds capture the true flux for all or most ecoregions. Also, an analysis of inversions for these cases that were performed with and without accounting for *a priori* flux temporal covariance (results not shown) confirms that accounting for this correlation *a priori* is necessary for recovering accurate uncertainty bounds at the monthly ecoregion scale. In contrast, the quality of inferred fluxes and uncertainties is lower for the F4d case regardless of spatial-temporal covariance assumptions, with only three of seven aggregated fluxes falling within two standard deviations of the true ecoregion flux in Figure 5.9. Overall, these results show that realistic *a posteriori* uncertainties can be recovered by the inversion at aggregated ecoregion scales, in addition to grid-scales (Table 5.5), as long as the diurnal cycle is estimated in the inferred fluxes, and both spatial and temporal correlation are considered in the *a priori* flux covariance matrix.

As expected, the addition of transport error degrades the quality of inferred ecoregion scale fluxes for all temporal resolutions, with the inferred values being closer to the mean continental flux as seen in Figure 5.9. The impact is greater for inversions that resolve the diurnal cycle (F3hr and F4d-diurnal), perhaps because these cases were originally better able to resolve ecoregion-scale variability in a setup without transport model error. In addition, for under-constrained areas (e.g. Tropics and Desert & Xeric Shrubland), random noise may obscure the diffuse signal from these areas more than in the near-field of the tower locations. Despite the fact that random transport errors are shown here to degrade inversion quality, uncertainty bounds for the cases resolving the diurnal cycle are still realistic for most eco-regions (results not shown). Again, it cannot be concluded from this analysis how systematic non-random transport errors would affect the inversion.

At the North American continental scale, inferred fluxes from all temporal resolutions, with and without transport error, are reasonably close to the true net flux (a difference of ≤ 0.25 μmol/(m^2s)). Consistent with previous results, the F4d case performs the least well, with the continental sink being significantly different from the true value. However, for all cases, as demonstrated by the narrower uncertainty bounds at the continental scale, results confirm that fluxes can be inferred more precisely at the continental scale than at smaller spatial scales.

Overall, the ecoregion-scale results confirm the importance of inferring the diurnal cycle directly in the estimated fluxes. This is particularly true in the near-field of the tower locations due to temporal aggregation errors, but it also appears to help constrain fluxes in the far-field as well. Random transport errors degrade the ecoregion-scale flux signal towards the continental mean, although their impact may potentially decrease as more areas become better-constrained by a growing measurement network.

5.4.5 Sensitivity tests with night-time data

In order to investigate the potential value of including night-time data for both Short and Tall/MBL towers, additional sensitivity tests were performed using the F4d-diurnal case. This resolution was chosen because of its comparable quality to the F3hr case with lower computational cost, which makes it most promising for annual or multi-year inversions. Inversions using reference covariance parameters and no transport error were used for these tests in order to isolate the impact of temporal aggregation errors associated with night-time measurements.

Results show that the continental grid-scale RMSE changes only marginally when night-time data are eliminated for the Tall/MBL towers (0.62 vs. 0.63 μmol/(m^2s)). However, including night-time data for both Short and Tall/MBL towers substantially increased the RMSE to 0.71 μmol/(m^2s), most likely due to the temporal aggregation errors associated with the smaller footprints for night-time measurements at the Short towers. An analysis of results at the aggregated ecoregion scale showed that the setup

used for most inversions in the current study, using night-time data only for the Tall/MBL towers, minimizes ecoregion scale RMSE's. This setup eliminates the temporal aggregation errors associated with including night-time data at the Short towers, while also allowing for a stronger constraint on far-field fluxes through night-time data from the Tall/ MBL towers. Again, in a real-data environment, the value of using night-time measurements from the Tall/MBL towers may be reduced for nights when the towers are within the PBL, and the PBL height is consistently over or under-estimated in the transport model.

5.5 Summary and conclusions

This study evaluated the constraint on CO_2 fluxes provided by atmospheric data from nine continuous measurement locations across the North American continent, within the context of a regional geostatistical inversion without any additional auxiliary variables. Estimating fluxes at a temporal resolution that can adjust the diurnal variability (F4d-diurnal and F3hr cases) was found to be crucial both for recovering covariance parameters directly from the atmospheric data, and for inferring ecoregion-scale fluxes that were statistically consistent with the "true" fluxes. Accounting *a priori* for both spatial and temporal covariance in the flux distribution was also found to be necessary for recovering accurate *a posteriori* uncertainty bounds on the estimated fluxes.

The poor performance of inversions that did not estimate the diurnal cycle (i.e. F4d) were due to the high temporal aggregation errors associated with not being able to adjust the strong diurnal and synoptic variability of the "true" fluxes, particularly near the measurement locations. For time periods outside of the growing season, the impact of temporal aggregation error may be lower, because fluxes are expected to be less variable, although the ability to infer accurate fluxes during the growing season is necessary for inferring accurate annual or multi-year carbon budgets. Also, while temporal aggregation errors may be of particular concern for geostatistical inversions because they do not assume the shape of the diurnal cycle *a priori,* any errors in the

diurnal cycle in prior flux estimates in synthesis Bayesian inversions would also yield temporal aggregation errors. This is likely to be of at least some concern, given the differences in the diurnal cycles predicted by different biospheric models.

In terms of the two flux resolutions resolving the diurnal cycle (i.e. F3hr and F4d-diurnal), both cases yielded flux estimates of comparable quality, despite the fact that covariance parameters estimated with the atmospheric data were more reasonable for the F3hr relative to the F4d-diurnal case. In fact, at the ecoregion scale for both of these flux resolutions (with no transport model error), the two standard deviation uncertainty bounds captured the true flux for all or almost all eco-regions. This is an encouraging result, showing that atmospheric data from only nine measurement towers sparsely located across the continent can be used to constrain ecoregion-scale fluxes in an idealized inversion setup without any additional auxiliary information from remote-sensing datasets or biospheric models, as long as the diurnal cycle is resolved in the fluxes. Between the two cases, the F4d-diurnal case has the additional advantage of estimating four times fewer fluxes than the F3hr case, yielding substantial computational savings that will become important for longer-term inversions. Therefore, given the comparable quality of inversion results for these two cases, performing regional inversions in a way that resolves a multi-day average diurnal cycle appears to be the most promising avenue for future inversions using real atmospheric data.

In this study, simulated random transport error was shown to decrease the quality of flux estimates in under-constrained areas at the ecoregion scale. It is important to note that non-random transport errors due to phasing or systematic biases, that may be more typical of the real inaccuracies in existing atmospheric transport models, may have more of an impact on flux estimation in the near-field where small errors in the transport matrices ($\mathbf{H}$) could translate into larger differences in inferred fluxes. The impact of transport errors on regional inversions has been investigated more thoroughly in other studies, e.g. Law et al. (2003), and should also be explored further in future studies focusing on North America.

Finally, it is important to note that real-data inversions are subject to additional complications in flux interpretation relative to synthetic data studies, due to potential biases introduced by the boundary conditions, non-random transport errors, and aggregation errors from flux variability at scales finer than the estimation grid-scale. Overall though, synthetic data experiments provide a baseline for the best achievable performance of real-data inversions, and help to highlight the impact of setup choices that could be obscured by the additional complexity associated with using real data.

In summary, synthetic data experiments were shown in this work to help illuminate the constraint on fluxes achieved by various regional inversion setup choices. The results suggest that even a fairly sparse network of continuous CO_2 measurements, used with no auxiliary information or prior estimates of flux variability in time or space, can be used to infer accurate monthly ecoregion scale CO_2 surface fluxes over North America, as long as the diurnal cycle is resolved in the estimated fluxes and both spatial and temporal flux covariances are accounted for *a priori*. Statistically significant sinks can also be recovered at the grid-scale, although uncertainties remain high at this fine spatial scale. The incorporation of additional atmospheric data and auxiliary variables in future real data geostatistical inversions can only help to further improve the recovery of continental CO_2 fluxes at fine spatial resolutions.

CHAPTER 6

Results from a regional grid-scale North American atmospheric CO_2 inversion for 2004 with a comparison to independent bottom-up flux estimates

6.1 Introduction

Carbon cycle scientists have been increasingly called upon to provide regional carbon budgets for helping to monitor and manage the anthropogenic component of the total CO_2 land-atmosphere exchange, and also to predict the future of the carbon cycle given a changing climate and human land-use choices. Atmospheric CO_2 inverse models can make a contribution toward these goals by taking advantage of the information in measured atmospheric CO_2 concentrations to infer the spatial distribution and rate of surface CO_2 flux occurring at relatively large spatial scales within a robust statistical framework.

Global inverse models using flask samples from approximately 80 locations worldwide in the NOAA-ESRL Cooperative Air Sampling Network, coupled with a global atmospheric transport model, can be used to infer continental-scale fluxes (e.g. Gurney et al., 2003; Baker et al., 2006; Rodenbeck et al., 2003; Mueller et al., 2008; Gourdji et al., 2008, or Chapter 4). However, the sampling of well-mixed air from marine boundary layer and high elevation locations in this measurement network limits the ability of such inversions to infer fluxes at finer spatial scales, particularly in regions with strong and highly variable CO_2 fluxes. The ability to use measured atmospheric CO_2 variability to resolve finer sub-continental-scale fluxes is critical for carbon budgeting at the scale of

political entities, e.g. states and provinces, as well as for helping to improve mechanistic model formulation to aid in future CO_2 flux prediction. In fact, the ability to provide consistent flux estimates for past years from atmospheric inversions and mechanistic models at a common scale remains a key challenge in carbon cycle science (Ciais et al., 2010).

Mechanistic, or bottom-up forward models scale up process-based understanding of photosynthesis and respiration from plot-level (e.g. Norby et al., 2010) and flux-tower studies (e.g. Baldocchi, 2008), while fossil fuel inventories rely on fuel sales, census data and air pollution measurements to estimate emissions at regional scales (Gurney et al., 2009). In contrast, atmospheric inversions (or top-down approaches) provide reasonably good constraints on carbon exchange at large scales, while uncertainties increase as the scale of estimation becomes finer. Theoretically, CO_2 flux estimates from bottom-up and top-down models should converge at intermediate regional scales (~500 km x 500 km), although the currently large spread among model estimates of CO_2 flux at such scales proves that this goal has not yet been attained (e.g. http://nacp.ornl.gov/mast-dc/int_synth_contreg.shtml).

In general, top-down CO_2 inverse models are limited by atmospheric data coverage (both in time and space), uncertainties in the atmospheric transport models used to relate concentration measurements to surface fluxes, and imperfections in the inversion setup for appropriately extracting information from the atmospheric measurements. However, an expanding measurement network, continuous data collection at an increasing number of sites, improvements in the quality of atmospheric transport models and improvements in inversion setup should all help to improve the ability of top-down models to infer fluxes at sub-continental scales with reasonable uncertainties. The atmospheric constraint could thereby be used to provide insight into the current spread of flux estimates across bottom-up models, and help improve the formulation of processes in these models for future flux prediction. In fact, atmospheric measurements of CO_2, if used appropriately, provide the main hope for validating

mechanistic models given that there is no way to directly measure regional-scale carbon exchange.

Regional inversions that estimate fluxes on a fine grid (e.g. 1°x1°) for a limited domain have emerged as a means to take advantage of continuous data collection at continental, low-altitude sites in a computationally feasible setup. This fine grid can help to reduce aggregation errors (Kaminski et al., 2001; Engelen et al., 2002) associated with estimating coarser-scale fluxes with highly variable measurement data. In addition, the limited domain makes it possible to take advantage of computationally expensive Lagrangian transport models that can adequately resolve the near-field around the measurement locations (Gerbig et al., 2008; Lin et al., 2003). However, regional inversions bring additional complications relative to global inversions, given that there is a need to specify atmospheric CO_2 boundary conditions for the region of interest, which have been show to have a large influence on the resulting flux estimates (Peylin et al., 2005; Göckede et al., 2010b). Also, the appropriate use of continuous, continental measurement data in an inversion presents a challenging problem, given the combined influence of the diurnal cycle of biospheric fluxes, heterogeneous land cover, point source fossil fuel emissions and complex atmospheric transport on the atmospheric CO_2 concentrations.

This chapter presents a regional grid-scale geostatistical inversion (Michalak et al., 2004; Mueller et al., 2008; Gourdji et al., 2008, or Chapter 4; Gourdji et al., 2010, or Chapter 5) for 2004 over North America using a limited sampling network of 9 towers collecting continuous CO_2 measurements on the continent in this year, as well as available flask and aircraft data. This inversion approach estimates fluxes at finer scales in both space and time than other published inversion studies for the same domain (e.g. Peters et al., 2007; Deng et al., 2007; Schuh et al., 2010; Butler et al., 2010), thereby reducing the impact of aggregation errors. This is an advantage of the current inversion, given that these errors have specifically been shown to have a negative impact on flux estimation for the North American continent using continental in-situ measurements (Gerbig et al., 2003b; Schuh et al., 2009; Gourdji et al., 2010, or Chapter 5).

In addition, the geostatistical inversions presented here are designed to help disentangle the relative influence of the atmospheric data constraint and the bottom-up prior flux estimates used in other inversion approaches, allowing for a relatively independent comparison to bottom-up mechanistic model output. While a geostatistical inversion can be run without the inclusion of any process-based datasets, thereby providing a completely independent comparison to biospheric models, such datasets defined at the grid-scale can also be incorporated into the inversion in a manner analogous to multi-linear regression. These datasets can include fossil fuel inventories, remote sensing vegetative indices, climatological model output, or even biospheric model estimates of CO_2 flux. Statistical variable selection techniques are employed to select those variables that can optimally explain the signal in the atmospheric data, and the drift coefficients on these auxiliary variables are estimated simultaneously with the fluxes using the atmospheric data. These inferred coefficients can thereby help to provide insight into flux drivers and their relationship to CO_2 flux at the scale of estimation, while also confirming that the inversion setup is recovering fluxes that are consistent with process-based understanding. Finally, incorporating these datasets can help to downscale and extrapolate the signal contained in the atmospheric data to under-constrained areas.

In the current study, we present two main sets of inversion results. First, we run an inversion incorporating only the atmospheric data and a fossil fuel inventory dataset. The second presented inversion incorporates diurnally-varying datasets associated with the biospheric signal from the North American Regional Reanalysis (Mesinger et al., 2006). After isolating the biospheric component in the *a posteriori* total flux estimates from both inversions, we compare results with a collection of bottom-up flux estimates from 16 forward models that participated in the North American Carbon Program Regional Interim Synthesis (Huntzinger et al., in prep).

In this inter-comparison, three main features of the model estimates are compared: the grid-scale spatial patterns in different seasons (Section 6.4.1), the shape of the seasonal cycle at aggregated spatial scales (Section 6.4.2), and the total annual

carbon budget at both the grid-scale (Section 6.5.1) and two aggregated spatial scales (Section 6.5.2). Overall, these analyses of CO_2 flux across different scales help to provide insight into the strengths and weaknesses of the inversion setup and data constraint for 2004, as well as into the large spread of estimates from the forward models. Comparisons to other inversion studies are also included where appropriate.

6.2 Data and methods

The geostatistical inversion implemented here uses a setup optimized as part of a one-month pseudo-data study with a similar tower configuration for June 2004 (i.e. Gourdji et al., 2010 or Chapter 5). Relative to synthesis Bayesian inversion studies (e.g. Baker et al., 2006; Peters et al., 2007; Butler et al., 2010) that rely on both prior estimates of flux and their uncertainties, the geostatistical inversion approach relies more directly on the atmospheric data to estimate fluxes at fine spatiotemporal scales, optimize covariance parameters and incorporate fine-scale flux variability provided by process-based datasets. A brief review of the setup and methods implemented in this study is provided below; additional details are provided in Chapters 3 and 5.

6.2.1 Flux domain and resolution

This study estimates fluxes at finer spatial and temporal scales than previous inversion studies estimating fluxes over all of North America (e.g. Peters et al., 2007; Deng et al., 2007; Butler et al., 2010). By estimating fluxes directly at the grid-scale (e.g. 1°x1° in this study), rather than for eco-regions or large blocks covering several states or provinces, spatial aggregation errors are minimized in the solution. These errors, which are caused by incorrectly specified flux patterns within large regions, are particularly problematic when making use of measurements collected in areas with high flux variability (Schuh et al., 2009).

Temporal aggregation errors are also a concern when using continuous, continental measurement data. The analyses in Chapter 5 (Gourdji et al., 2010) showed the importance of explicitly estimating the diurnal cycle of fluxes when using continental

measurements over North America in a regional inversion, in order to avoid biasing larger-scale post-aggregated fluxes. In contrast to geostatistical inversions, most synthesis Bayesian inversions pre-subtract the influence of the biospheric diurnal cycle on measurement data *a priori*. However, the large magnitude of the bias in inferred fluxes seen in Chapter 5 may also imply bias in other synthesis Bayesian inversions due to errors in the shape, phase and amplitude of the diurnal cycle specified by the bottom-up prior estimates of Net Ecosystem Exchange (NEE; Huntzinger et al., 2011).

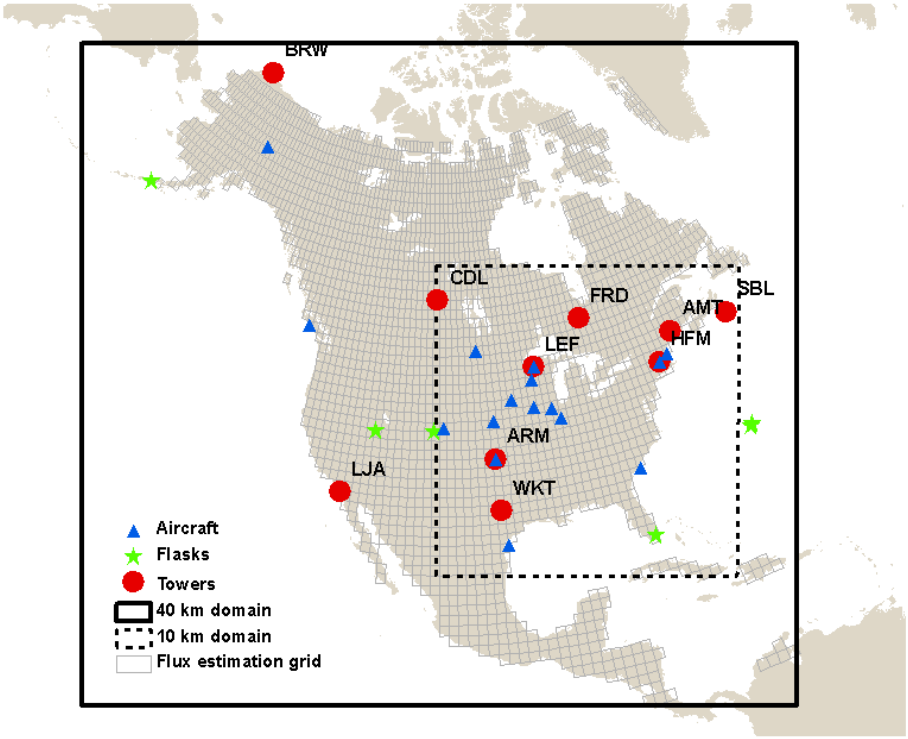

Figure 6.1: Domains of nested WRF winds, flux estimation grid, and the locations of towers, flask & aircraft measurements used in the inversions. The map is displayed using an equal-area projection.

Fluxes are estimated in this study from 10°N to 70°N and 50°W to 170°W, resulting in 2635 land grid-cells in the domain (Figure 6.1). Given the results of Chapter 5, inferred fluxes are allowed to vary at a 3-hourly timescale, resulting in approximately

8 million fluxes (2635*366*8) for the year. In order to save computational expense, resolving fluxes for a 4-day average diurnal cycle was also considered, but ultimately rejected due to significant decreases in the quality of the inferred flux estimates using this coarser temporal resolution.

6.2.2 Atmospheric transport, data and boundary conditions

6.2.2.1 Atmospheric transport model

Atmospheric transport models are used to provide a critical component of inversions, i.e. the footprints that link measured concentrations to upwind surface CO_2 fluxes. For this study, the Stochastic Time-Inverted Lagrangian Transport Model (STILT) was chosen, given that Lagrangian models allow for improved representation of the near-field around a measurement location relative to grid-scale Eulerian models (Lin et al., 2003), and are therefore ideal for taking advantage of continuous measurements sited in areas with high CO_2 flux variability and complex transport phenomena. STILT has also been used in other inversion studies estimating CO_2 fluxes at regional scales (Gerbig et al., 2003b; Lin et al., 2004; Matross et al., 2006; Trusilova et al., 2010).

STILT was used in conjunction with high-resolution winds from the Weather Research & Forecasting (WRF) model to simulate backward particle trajectories for each measurement location and time period. For this study, the WRF winds had a 10 km resolution for the eastern portion of the continent (capturing the near-field for 7 of the 9 towers), nested within a 40km resolution winds for the rest of the continent, as shown in Figure 6.1. The Grell-Devenyi cumulus parameterization was used for both the 10 and 40-km resolution winds (Grell & Devenyi, 2002). The ability to nest high-resolution wind fields for specific locations within WRF helps to capture small-scale dynamics near the receptor locations that can have a significant impact on simulated particle trajectories (Nehrkorn et al., 2010). WRF has specifically been shown to have good performance in simulating small-scale atmospheric CO_2 dynamics in the northeastern continental United

States (Nehrkorn et al., 2010), and also at coastal sites (Sarrat et al., 2007; Ahmadov et al., 2009).

Concentration footprints for the inversion were calculated within STILT by releasing 500 particles for each measurement location and collection time, and then tracing the particle trajectories backwards for 10 days using the WRF wind fields. The proportion of particles from each release reaching a given flux location and time interval was used to derive sensitivities to the measured concentration (Lin et al., 2003). The atmospheric transport model used in this study will henceforth be referred to as WRF-STILT.

6.2.2.2 Atmospheric CO_2 concentration measurements

This study takes advantage of continuous, calibrated CO_2 measurements taken at 9 observational locations unevenly spaced across the North American continent in 2004. These include two tall towers with a height of 457 and 396m, two marine boundary layer (MBL) towers less than 25m in height, and five other inland, continental towers ranging in height from 30 to 107m (Table 6.1, Figure 6.1). After preliminary data filtering to exclude low-quality flags and other obvious errors or anomalous spikes in the data, continuous measurements from all towers were averaged to a three-hourly timescale for use in the inversion. In addition, to maximize the atmospheric data constraint for this relatively data-poor year that precedes the recent expansion of the North American measurement network (Mueller et al., in prep.), available flask and aircraft measurements were included. Aircraft data from measurements below 4000m in altitude were used, where the strongest influence from surface CO_2 fluxes on the continent is expected (Gerbig et al., 2003a); flasks at coincident continuous tower locations were also excluded. Measurement locations and associated details for towers, flask and aircraft data are shown in Table 6.1 and Figure 6.1.

Table 6.1: Measurement locations, along with other identifying characteristics of the sites and data included in the inversion. The first 9 locations in the table have continuous data, while the last two represent flask and aircraft measurements from multiple locations across the continent. Figure 6.1 shows the locations of towers, as well as individual flask and aircraft measurement sites.

Measurement site code	**Site name**	**Site latitude/ longitude**	**Altitude above ground level (m)**	**Night-time data included?**	**Number of data-points for year**
LEF	Park Falls, Wisconsin	45.93N, 90.27W	396	1 and 4 am year-round	1500
WKT	Moody, Texas	31.32N, 97.33W	457	1 and 4 am year-round	959
SBL	Sable Island, Nova Scotia	43.93N, 60.02W	25	None	663
BRW	Barrow, Alaska	71.32N, 156.60W	10	None	248
ARM	Norman, Oklahoma	36.62N, 97.50W	60	None	879
HFM	Harvard Forest, Massachusetts	42.54N, 72.17W	30	None	558
AMT	Argyle, Maine	45.03N, 68.68W	107	None	795
FRD	Fraserdale, Ontario	49.84N, 81.52W	40	None	878
CDL	Candle Lake, Saskatchewan	53.99N, 105.12W	30	None	927
FLA	Flask samples from 6 sites	*See Figure 1*	0 to 4	None	153
AIR	Aircraft vertical profiles from 15 sites	*See Figure 1*	139 to 3999	None	943

With a perfect transport model, it would be possible to use all available CO_2 measurement data in an inversion to help improve the atmospheric constraint on flux estimates. However, systematic transport errors associated with certain measurement times and locations can lead to biased flux estimates, and potentially misleading scientific conclusions (Lin and Gerbig, 2005; Prather et al., 2008). Systematic transport model errors that affect inversions mostly relate to misrepresenting the height of the night-time Planetary Boundary Layer (PBL) (Denning et al., 2008), and also the height of the PBL during morning and evening transition times (A. Andrews, personal communication).

Therefore, this study takes a similar approach to other inversion studies (e.g. Peters et al., 2007; Schuh et al., 2010; Göckede et al., 2010a) in relying primarily on afternoon measurements of CO_2 concentration when vertical convective mixing is strongest, and the height of the PBL is presumably well-modeled within WRF-STILT. Specifically, 3-hourly averages centered at 1 and 4pm were included year-round for all towers. This study additionally included some morning and evening data (3-hourly averages centered at 10am and 7pm) during the height of the growing season when the air should also be well-mixed, with time periods selected for inclusion based on sunrise and sunset for each tower. Also, following the data choices for the CarbonTracker data assimilation system (Peters et al., 2010; Andrews et al., personal communication), some night-time data (3-hourly averages centered on 1 and 4am) were included for the two tall towers (LEF and WKT, Table 6.1). Finally, given presumed difficulties in appropriately modeling the coastal land-sea breeze, some of the flask and aircraft data collected on the Pacific coast were excluded, given a substantial misfit at these sites between measured concentrations and modeled fluxes transported forward to the measurement locations.

In addition to systematic transport model errors, using real CO_2 continuous measurement data in a regional inversion is subject to other concerns and limitations. First, measurement instruments are subject to failure, and therefore, there are several long gaps in the concentration record for certain towers. For example, at Harvard

Forest, 45% of the potential measurements are missing for the year, particularly in the early part of the year, while at Moody, Texas, measurements are missing for January, August and most of September. Other shorter data gaps occur throughout the year for all towers.

Second, some of the variability in the measurement data is due to very local influence (Gerbig et al., 2009) which cannot be resolved by the transport model, the driving meteorological data, or the flux estimation grid, leading to representation errors in the inversion. These data should ideally be excluded, although how to identify purely local influence remains a challenging research question. Therefore, with the exception of filtering out extreme fossil fuel spikes (> 30 ppm over background air), this study took the approach of not attempting to filter the data for local variability.

Third, measurements are influenced by both land and oceanic fluxes, although oceanic fluxes were not explicitly resolved in this study. This is an issue particularly for marine boundary layer locations where mixing with ocean air can dilute the influence of land fluxes, thereby leading to misleading land flux estimates if this is not taken into account in the inversion framework. Therefore, filtering was applied to exclude measurements with substantial oceanic influence (defined as less than 10% of the total sensitivity coming from land regions or a less than 0.5 ppm/($\mu mol*m^{-2}*s^{-1}$) integrated land footprint). After all data filters were applied, the number of included data points per tower, accounting for data gaps, is shown with other tower information in Table 6.1.

6.2.2.3 Continental boundary conditions

Regional inversions necessitate the use of boundary conditions that represent the influence of fluxes occurring outside the domain of interest, i.e. the North American land mass, on the measured concentrations. This influence must be subtracted from the measurements before they can be used in the inversion to infer CO_2 fluxes over North America. However, these boundary conditions also have a strong potential to affect the total continental carbon budget. Therefore, results are presented here using two plausible sets of CO_2 boundary conditions, one optimized as part of the

CarbonTracker global CO_2 data assimilation system (Peters et al., 2007), and the other derived more directly from measured CO_2 concentrations, specifically those taken from the GlobalView database in the Atlantic & Pacific Oceans (Andrews et al., in prep).

The CarbonTracker boundary conditions represent model output, and are subject to biases in the inferred global fluxes resulting from the data assimilation system. In fact, the GlobalView dataset was derived in response to known seasonal biases in the CarbonTracker CO_2 fields (Peters et al., 2010), and preliminary validation using measurements of "clean" ocean air sampled at Pacific and Atlantic coast locations shows that this new dataset helps to correct these biases (A. Andrews, personal communication). These corrections led to a systematic offset between the two datasets, with the GlobalView values on average about 0.5 ppm lower than the CarbonTracker values, although this offset tends to be higher during the growing season, and is slightly lower for measurement locations on the East coast. While the GlobalView boundary conditions are more empirical in that they rely on direct measurements of MBL air, they also lack synoptic variability as compared to the CarbonTracker CO_2 fields. Furthermore, they lack longitudinal variability, given that the GlobalView dataset is currently defined as a single vertical curtain representing an average of flask and aircraft-based measurements from the Atlantic and Pacific Oceans.

These two sets of boundary conditions are used in conjunction with WRF-STILT to quantify the influence of sources and sinks outside the domain on measured concentrations, and this impact is then pre-subtracted from the data before use in the inversion.

6.2.3 Inversion setup and algorithm

6.2.3.1 Geostatistical inversions

Geostatical inverse modeling, first developed in the field of groundwater hydrology (e.g. Hoeksema & Kitanidis, 1984; Zimmerman et al., 1998), has more recently been used in atmospheric applications to identify trace gas sources and sinks (Michalak et al., 2004; Mueller et al., 2008; Gourdji et al., 2008, or Chapter 4). Geostatistical

inversions have also been applied in a pseudo-data regional inversion setup for North America (Gourdji et al. 2010, or Chapter 5), where the setup emerging from that study is used to inform the current work.

Geostatistical inversions are similar to synthesis Bayesian inversions in many respects, except for two key differences. First, while Bayesian inversions start from a set of explicit prior flux estimates derived from biospheric models, fossil fuel inventories, fire emission estimates, and ocean ship-track data, geostatistical inversions have no explicit prior. Instead, the "prior" term in the geostatistical inversion objective function is replaced with a statistical linear trend whose coefficients are optimized using the atmospheric data as part of the inversion. This trend can be as simple as a single mean flux in both time and space, avoiding the use of any process-based information in the model altogether. Alternatively, fluxes can be modeled as having a trend with process-based datasets, in a statistically rigorous manner, to help inform flux estimates at fine scales and in under-constrained areas. By their ability to avoid the use of prior fluxes or incorporate them in a flexible way consistent with the atmospheric data constraint, geostatistical inversions can help to diagnose the inversion quality of synthesis Bayesian inversions, by decomposing the information provided by the prior fluxes vs. that contained in the atmospheric measurements.

Secondly, geostatistical inversions incorporate an *a priori* spatial and/or temporal covariance between estimated fluxes (or flux residuals), which helps to regularize the solution, and also makes it possible to resolve fluxes at fine scales necessary for avoiding aggregation errors. Spatial covariance assumptions are not unique to geostatistical inversions, and in fact have been introduced into other grid-scale inversion studies as well (e.g. Rödenbeck et al., 2003; Carouge et al., 2010a,b; Chevallier et al., 2010). However, these previous studies estimated fluxes on a coarser grid, and/ or had more limited freedom to adjust fluxes in time as compared to the current work. Also, these studies specified the parameters of the covariance matrices based on an analysis of variability in measured data or bottom-up flux estimates (e.g. Baker et al., 2006), whereas these parameters are estimated in the current study

directly using the signal in the atmospheric measurements, as described further in Section 6.2.3.2.

The geostatistical inversion objective function was minimized, and fluxes ($\hat{\mathbf{s}}$) and their uncertainties ($V_{\hat{\mathbf{s}}}$), as well as drift coefficients ($\widehat{\boldsymbol{\beta}}$) and their uncertainties ($V_{\widehat{\boldsymbol{\beta}}}$), were estimated using equations 3.2 through 3.7. Please refer to Chapter 3 for more details on the implementation of the geostatistical inverse modeling approach.

6.2.3.2 Covariance matrix structure and parameter optimization

The $\mathbf{Q}$ matrix describes the spatiotemporal correlation structure of flux residuals that cannot be explained by the ancillary environmental variables included in the linear model of the trend. Two parameters, the variance (σ^2), and the spatial correlation length parameter (l), are used to populate the $\mathbf{Q}$ matrix assuming an exponential decay, as shown in equation 3.9.

Flux covariance parameters are allowed to vary at a monthly timescale in the current study, given that these parameters have been found to have a strong seasonal cycle (Huntzinger et al., 2010) that should be appropriately accounted for in order to yield realistic flux estimates from the inversion. *A priori* temporal correlation was additionally considered, as in Chapter 5 (i.e. Gourdji et al., 2010), but was ultimately rejected due to unrealistic extrapolation of flux patterns in both space and time in the current work.

However, in Chapter 5, *a priori* temporal covariance assumptions were shown to help recover more realistic flux uncertainties from the inversion at temporally aggregated scales. In this study, the estimated uncertainties recovered with a real data setup and a 3-hourly flux resolution were also seen to be artificially low at the post-aggregated monthly and annual timescales. Therefore, the calculated uncertainties that correspond with the presented fluxes are not shown in this study. Instead, those corresponding to an inversion with a 4-day average diurnal cycle flux resolution are shown instead to give a general idea of the atmospheric data constraint provided by the inversion in different areas of the continent. This coarser-scale flux resolution was seen

to have less of an impact on uncertainty collapse at aggregated temporal scales. The recovery of realistic confidence intervals from an inversion, particularly at the annual timescale, remains an area of active research (e.g. Peters et al., 2010).

The model-data mismatch covariance matrix ($\mathbf{R}$) describes how well the "true" flux solution should be able to match the recorded measurements, given errors associated with atmospheric transport, the measurement instruments, and the coarse grid of the inversion and transport model relative to the scales of variability in the true fluxes (e.g. Kaminski et al., 2001; Engelen et al., 2002). In the current study, the model-data mismatch matrix ($\mathbf{R}$) remains diagonal as in previous inversion studies, although as in Chapter 5, a different model-data mismatch was optimized per tower, with separate additional values for flask and aircraft data. This allows the inversion to de-weight continuous measurement locations where the transport model may be of poorer quality or local variability in the measurement data is difficult to resolve, although the model-data mismatch values for flask and aircraft measurements represent an average across sites. In addition, the optimized model-data mismatch values were allowed to change monthly in order to account for seasonal variations in the quality of the transport model and inversion setup.

The covariance parameters used to construct the $\mathbf{R}$ and $\mathbf{Q}$ covariance matrices are optimized here using the Restricted Maximum Likelihood method (e.g. Kitanidis, 1995) with the atmospheric measurements, minimizing equation 3.12. (Please see Chapter 3, Section 3.2 for more information about this algorithm.) Chapter 5 showed that not only was it possible to estimate reasonable covariance parameters using the atmospheric measurement data, but that fluxes estimated using these parameters were of comparable or superior quality to those estimated using the "true" covariance parameters associated with the underlying fluxes. $\mathbf{R}$ and $\mathbf{Q}$ parameters are optimized simultaneously in the current study, such that inversion artifacts (such as dipoles, or artificially strong counteracting sources & sinks) are minimized in the inferred flux distribution.

6.2.3.3 Auxiliary variable selection

While a geostatistical inversion can estimate fluxes with a very simple model of the trend (e.g. monthly means over the land and oceans, as in Mueller et al., 2008), grid-scale environmental datasets with a mechanistic relationship to CO_2 flux also have the potential to improve estimates if incorporated into the inversion in a statistically rigorous manner (e.g. Gourdji et al., 2008, or Chapter 4). First, these variables help to downscale the atmospheric signal, particularly in areas well-constrained by the measurements. Second, the selected variables and their inferred relationship to flux ($\widehat{\boldsymbol{\beta}}$) can provide process-based understanding of flux drivers at the scale of the inversion, while also helping to validate the inversion model setup if the selected variables and $\widehat{\boldsymbol{\beta}}$ values are consistent with process-based understanding. Finally, auxiliary variables extrapolate the atmospheric signal to under-constrained areas using process-based relationships derived from the CO_2 concentration measurements, although the selected variables and inferred drift coefficients ($\widehat{\boldsymbol{\beta}}$) may be partly biased by the agricultural and forested regions in the center of the continent sampled by the 2004 measurement network. For example, under-sampled arid regions of the continent may have additional significant flux drivers, and/ or somewhat different relationships between driving variables and CO_2 flux.

For the first inversion presented in this paper, we include only a single fossil fuel inventory dataset in the trend ($\mathbf{X}$). In the continental United States, this dataset consists of diurnally and seasonally varying estimates from version 1.4 of the Vulcan database for 2002 (Gurney et al., 2009), scaled up to 2004 total emissions. In Central America, Mexico and Canada, estimates were included from a monthly-varying dataset merging information from British Petroleum fuel statistics, remotely-sensed night lights, and the existing Carbon Dioxide Information Analysis Center (CDIAC) fossil fuel emission inventory. This latter dataset was recently developed for global inversions using GoSat satellite data (Oda and Maksyutov et al., 2010). By including only this single combined continental fossil fuel inventory in $\mathbf{X}$, we help to localize the spatial patterns of the fossil fuel emissions within the total CO_2 flux, while also estimating biospheric fluxes from the

inversion that are completely independent of mechanistic, forward model output. This inversion will henceforth be referred to as the "FF-only" inversion.

For the second presented inversion in this study, we incorporate auxiliary environmental variables associated with the biospheric signal into the model of the trend. Overall, eleven 3-hourly datasets from the North American Regional Reanalysis (NARR; Mesinger et al., 2006) were considered, as well as two derived precipitation variables (16 and 30-day lagged average precipitation). A subset of these 13 variables (Table 2), which are defined for all flux locations and time periods, was chosen for inclusion in a statistically rigorous manner using the Bayes Information Criterion (BIC) (Schwarz, 1978) combined with the Branch-and-Bound algorithm (Yadav et al., in prep.), as described in more detail in Chapter 3.

The NARR datasets provide information regarding the diurnal cycle of water availability and radiation, as well as derived quantities such as evapotranspiration and canopy conductance, derived from the Noah Land Surface Model (Ek et al., 2003) within NARR, which should have more of a linear relationship to CO_2 fluxes. While remote-sensing datasets, such as Leaf Area Index or Fraction of Photosynthetically Active Radiation from the MODIS instruments (e.g. Yang et al., 2006), could also provide useful information regarding the seasonal cycle and spatial distribution of CO_2 flux, they are defined only at a weekly timescale, which complicates their correlation to diurnally-varying fluxes. Also, the evapotranspiration and canopy conductance variables from NARR implicitly include the Normalized Difference Vegetation Index from the AVHRR instrument, thereby providing an alternative to MODIS datasets.

For the inversion including NARR datasets (termed the "NARR" inversion), we pre-subtract the influence of the fossil fuel inventory on the measurements before performing variable selection and the inversion itself. While errors in the fossil fuel inventories will thereby become aliased onto the inferred biospheric fluxes with this setup, initial tests showed a recovered drift coefficient on the fossil fuel inventory near one (see Section 6.3.3), implying that the inventory dataset used in this study was a reasonable approximation. By pre-subtracting the fossil fuel influence from the

measurements, we also reduce potential covariance in the inferred drift coefficients between the fossil fuel inventory and biospheric datasets due to covariance in the underlying processes, e.g. re-growing forests and high emissions in the eastern continental United States, or reduced populations and industrial activity in arid and snow-covered areas. Such covariance would confound flux interpretation by making it difficult to separate the biospheric and fossil fuel signals *a posteriori*.

6.2.4 Evaluation of inferred fluxes

Given that there is no way to directly measure CO_2 flux at the grid-scale (approximately 100 km x 100 km), nor even at coarser regional or continental scales where atmospheric inversions perform best, validating inversion flux estimates remains a challenge. Some studies have taken the approach of excluding some atmospheric measurements from use in the inversion (e.g. aircraft measurements), and then comparing concentrations resulting from inferred fluxes to these excluded observations to evaluate the inversion (e.g. Peters et al., 2007; Chevallier et al., 2010). However, given the limited amount of atmospheric CO_2 measurement data in 2004, we wanted to use all available data-streams to help improve flux estimates and fill in regions of the continent otherwise under-constrained by the nine tower network. Also, by comparing transported fluxes to measurement data excluded from the inversion, it is difficult to interpret the relative impact of biases in the inversion fluxes vs. errors in the transport model, limiting the power of such a technique.

Rather, this study relies on pseudo-data inversions, as in Chapter 5, and the inter-comparison to the forward models to assess inversion biases and limitations. In a pseudo-data inversion, synthetic measurements are created from a set of bottom-up flux estimates, thereby providing a "true" solution for comparison to inferred fluxes at various spatial and temporal scales. Pseudo-data test results are not explicitly shown here, although they were used to guide the presentation and interpretation of the real data results, and are discussed at relevant points in the text. Also, the inter-comparison with the forward model estimates of NEE, as well as the auxiliary variable selection

results from among the NARR datasets, both help to sense-check that inversion results are consistent with process-based understanding of biospheric CO_2 flux. While it is not possible to remove all possible sources of bias in inversions, particularly systematic transport model errors, the inversions presented here were deemed sufficiently representative enough of the atmospheric data constraint to begin to give insight into the spread of the forward models included in the NACP Regional Interim Synthesis.

6.3 Results: Footprints, covariance parameters and auxiliary variables

6.3.1 Concentration footprint analysis

In addition to constraining fluxes in the inversion, the concentration footprints derived from WRF-STILT allow one to assess which portions of the continent are "seen" by a certain set of measurements. Given the limited network collecting continuous CO_2 measurements in 2004, not all portions of North America are equally well-constrained, as can be seen in the yearly-average footprint shown in Figure 6.2. Not surprisingly, the best-constrained areas are in the central and eastern continental United States, and a large part of Canada near the measurement locations. The under-constrained areas are in the tropics, northwest Canada and Alaska. The areas with a partial constraint are in the western and southeastern continental United States. The contour line in this figure identifies a high sensitivity area where fluxes are reasonably well-constrained throughout the year. Pseudo-data testing revealed that annual budgets could be most reliably interpreted in this area, as will be discussed further in Section 6.5.2.

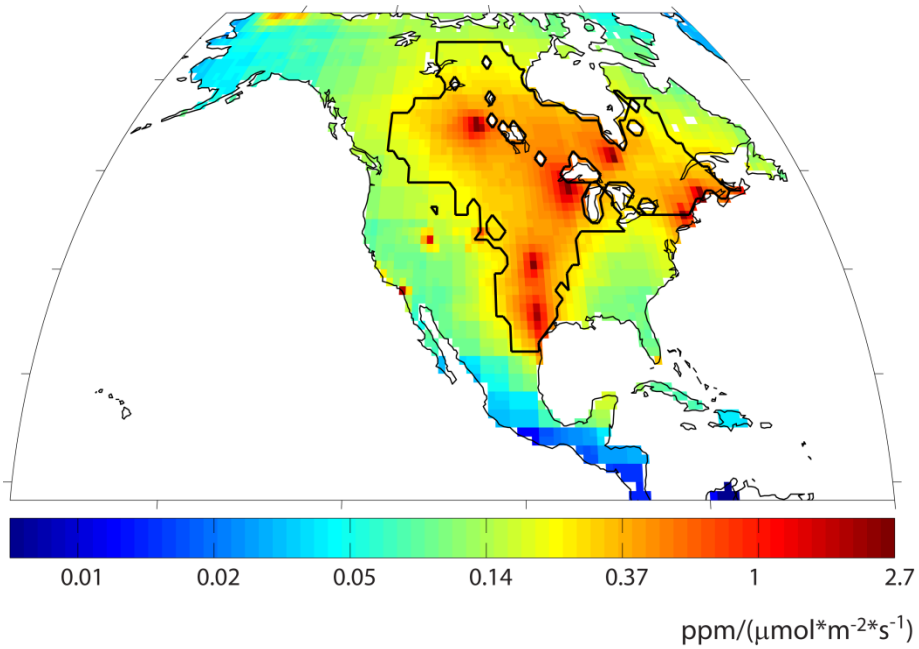

Figure 6.2: Yearly average integrated footprint for all measurement locations. The contour line shows the high sensitivity areas used for flux interpretation at the spatially aggregated annual scale (Figure 6.8).

6.3.2 Inferred covariance parameters

The monthly flux covariance parameters inferred using the atmospheric data with the RML method provide insights into the underlying variability of the true flux field, and how this changes throughout the year. The inferred monthly model-data mismatch parameters by measurement location are representative of the ability of the inversion setup and transport model to appropriately take advantage of surface flux information contained in the measured CO_2 concentrations.

Figure 6.3a shows inferred monthly spatial flux covariance parameters from the FF-only inversion, where this covariance is among flux residuals from the trend, which represent the total biospheric flux in this case. These monthly covariance parameters show realistic seasonality also seen in biospheric model output (Huntzinger et al., 2010), with the most variable fluxes during July and August (highest variance and shortest correlation length), and the least variable in the dormant season from November through April (lowest variance and longest correlation length). The inferred correlation lengths (i.e. 3*l*, see Chapter 3, Section 3.2) range from ~2000 km in September to ~8500

km in January, somewhat shorter than that seen in modeled biospheric flux estimates (Huntzinger et al., 2010). In the NARR inversion, inferred flux variances and correlation lengths are reduced relative to the FF-only inversion, by about ~13 and ~33% respectively, showing that the NARR variables are able to explain some, but not all, of the coherent variability in the inferred flux distribution.

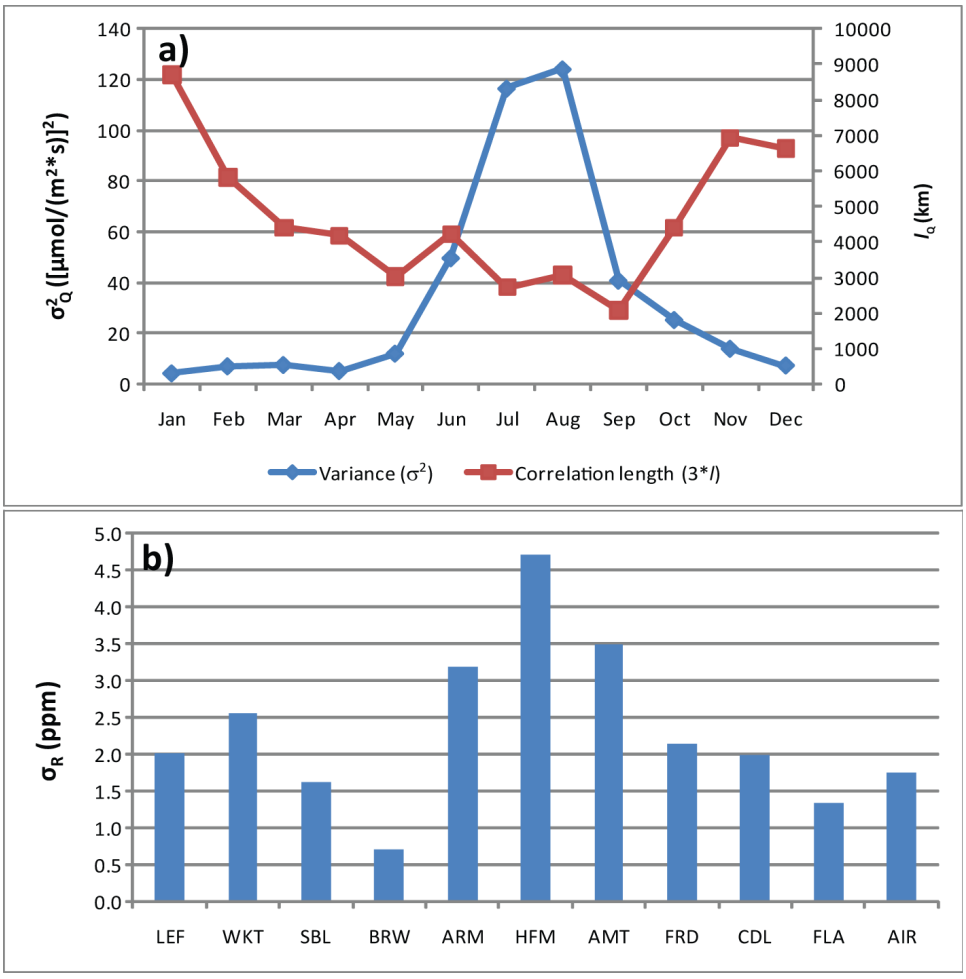

Figure 6.3: Optimized covariance parameters using the RML algorithm with the atmospheric measurements, for the FF-only inversion with GlobalView boundary conditions. a) Monthly flux covariance parameters, i.e. the sill variance (σ^2_Q) and correlation length parameter (l). b) Square root of the averaged monthly model-data mismatch variances (σ_R), (weighted by the number of data-points in each month), for 9 towers, flask and aircraft data.

A yearly average of the monthly model-data mismatch variances for the 9 measurement locations and flask and aircraft data is shown in Figure 6.3b, where these averages are weighted by the number of data points associated with each estimated parameter (per measurement site and month). This figure shows that the tower with the highest model-data mismatch is Harvard Forest. This tower is sited in a forested area about 100 km west of Boston, and even closer to Worcester and Springfield, Massachusetts, while the 1°x1° gridcell containing this site includes several other small towns and developed areas. Therefore, the difficulty in matching the data at this tower is most likely due to spatial aggregation and representation errors from fossil fuel plumes and heterogeneous land cover in the region. The two towers with the lowest model-data mismatch are SBL and BRW, the two MBL sites sampling relatively well-mixed air and/ or low flux variability areas. Similarly, the flask measurements, collected in either MBL or high-altitude locations, have low model-data mismatch. Interestingly, FRD and CDL, which are short towers in a forested and agricultural region respectively, have lower model-data mismatch than the three other short towers. Given their northern location in the domain where the horizontal grid-cell size is smaller (see Figure 1), this may suggest potential improvements in inversion performance associated with resolving fluxes at finer spatial scales, perhaps the resolution of the driving winds in the transport model (in this case a 40-km grid for all North America).

Model-data mismatch variances can also vary significantly within the year (results not shown). For example, the optimized model-data mismatch at ARM is 1.1 ppm in February, 6.1 ppm in April and 2.4 ppm in July. The high model-data mismatch variance in April for ARM implies that the strength of local uptake visible in the measurement data in this month, perhaps due to the spring wheat crop planted in the near-vicinity of the tower, is not being interpreted to the full extent by the inversion. Again, a finer spatial flux resolution could potentially help to more appropriately use the information contained in this data in future inversions.

6.3.3 Auxiliary variable selection and inferred drift coefficients ($\widehat{\boldsymbol{\beta}}$)

Introducing covariates associated with the biospheric signal into the $\mathbf{X}$ matrix gives an opportunity to identify significant flux drivers, and infer the relationship between these variables and flux as seen by the atmospheric data at the resolution of the estimated fluxes. Also, if the selected variables and inferred relationships between the datasets and CO_2 flux (i.e. $\widehat{\boldsymbol{\beta}}$) are consistent with process-based understanding, this helps to validate the inversion setup. Table 6.2 shows the selected variables and inferred $\widehat{\boldsymbol{\beta}}$ values for the FF-only and NARR inversions.

Given that fossil fuel emissions are relatively well-known in comparison to the biospheric signal, the inferred drift coefficient on this dataset should be close to one, as seen in pseudo-data tests with perfect transport. Values other than one could imply problems with the inversion setup, systematic transport model errors, or errors in the spatial and temporal patterns of emissions in the inventory dataset. The results in Table 6.2 are encouraging in that the inversion using GlobalView boundary conditions infers a $\widehat{\boldsymbol{\beta}}$ on the inventory of 1.00, and the value inferred using the CarbonTracker boundary conditions is 0.98. While it is impossible to know exactly what is driving these values (for example, an oversampling of the morning and evening rush hours could compensate for a less than perfect correlation with the inventory dataset), these results suggest that the data choices made for the presented inversion may be helping to minimize the impact of systematic transport model errors on the estimated fluxes. The $\widehat{\boldsymbol{\beta}}$ near one also makes it easier to separate the biospheric and fossil fuel contributions to the total flux *a posteriori*, particularly for the grid-scale spatial patterns, although again some aliasing of one signal onto the other is possible due to imperfections in the inventory dataset and inversion setup.

The NARR variables selected for inclusion in the inversion are shown in Table 6.2, and are very consistent with process-based understanding of CO_2 flux, providing additional support for the inversion setup implemented in this study. For example, evapotranspiration is most strongly correlated with uptake (as indicated by a negative $\widehat{\boldsymbol{\beta}}$), which is consistent with the near linear relationship between this variable and flux as

described in mechanistic theories of plant physiology (Bonan, 2008) and also found in other statistical studies of CO_2 flux using eddy-covariance measurements (e.g. Mueller et al., 2010; Yadav et al., 2010).

*Table 6.2: Selected variables and associated $\widehat{\boldsymbol{\beta}}$ values from the FF-only and NARR inversions, using both sets of boundary conditions. The fossil fuel inventory is in flux units (μmol/m^2*s), such that the $\widehat{\boldsymbol{\beta}}$ represents a scaling factor on this dataset. Auxiliary variables from NARR were normalized to have zero mean and variance of one, such that $\widehat{\boldsymbol{\beta}}$'s are directly comparable. Shaded cells in the table represent variables not included in the model selection for a given trend, whereas variables with dashes ('---') were considered but not selected by the BIC/ Branch & Bound algorithm. All $\widehat{\boldsymbol{\beta}}$ values were significantly different from zero at $2\sigma_{\widehat{\boldsymbol{\beta}}}$.*

	Trend with fossil fuels only		***Trend with NARR variables***	
	GlobalView BC's	CarbonTracker BC's	GlobalView BC's	CarbonTracker BC's
Fossil Fuels	1.00	0.98		
Canopy Conductance			---	---
Downward Shortwave Radiation			---	---
Evapotranspiration			-1.54	-1.60
Precipitation Rate			0.24	0.23
Relative Humidity			---	---
Specific Humidity			0.10	0.30
Soil Moisture			---	---
Air Temperature (@ 2m)			---	-0.21
Plant Canopy Water Content			---	---
Snow depth			---	---
Snow cover (%)			-0.13	-0.19
16-day Lagged Precipitation			---	---
30-day Lagged Precipitation			0.18	0.14

The positive $\widehat{\boldsymbol{\beta}}$ values associated with precipitation rate, 30-day lagged precipitation and specific humidity are consistent with process-based understanding of the drivers of heterotrophic respiration (Ise and Moorcroft, 2006). The source associated with precipitation rate at a 3-hourly timescale is consistent with flux tower studies showing pulses of respiration following rain events (Baldocchi, 2008), while the 30-day lagged precipitation helps to better explain respiration fluxes associated with longer-term soil moisture and water availability. Specific humidity, or the mass of water vapor per unit mass of air, scales with both water availability and temperature, and can therefore additionally help to explain the well-known temperature dependence of respiration (Lloyd & Taylor, 1994). Snow cover acts to reduce respiration sources, which is consistent with process-based studies showing that snow can act to trap soil respiration fluxes until the spring thaw (Kelley et al., 1968; Björkman et al., 2010). The inferred $\widehat{\boldsymbol{\beta}}$ values using the two sets of boundary conditions are mostly consistent, although air temperature is additionally selected using the CarbonTracker dataset. An analysis of the *a posteriori* covariance between the $\widehat{\boldsymbol{\beta}}$ values (i.e. $\mathbf{V}_{\widehat{\boldsymbol{\beta}}}$) shows that air temperature in this trend is mainly helping to correct the signal associated with specific humidity and percent snow cover.

It should be noted here that not all processes that affect CO_2 flux can be included in a statistical model like this one, especially discrete events like the escape of trapped respiration fluxes directly following snow melt. Also, this study only chose to examine variables available from the NARR, and excluded other possible datasets that could help to explain fluxes (e.g. a fire emission inventory). Finally, while evapotranspiration and canopy conductance implicitly incorporate a measure of biomass (specifically the Normalized Difference Vegetation Index), the trend with NARR variables still lacks a measure of substrate availability for respiration. Despite these concerns as well as potential biases in the NARR datasets themselves, the design of the inversion is such that variability excluded or incorrectly specified in the linear model of the trend are still included (or corrected) in the best estimates of flux through the spatially-correlated stochastic component of the best estimate (see eq. 3.7).

6.4 Results: Comparison of biospheric flux estimates to forward models and other inversions

Given that there is no direct validation data for regional-scale CO_2 land-atmosphere exchange, an inter-comparison of inversion results with forward model estimates is bound to be somewhat inconclusive as to which models are more or less "correct." However, because of this lack of validation data, it is important to be able to identify strengths and weaknesses in specific models by a careful inter-comparison of different modeling approaches. Such a comparison may point towards an approach for obtaining improved estimates in future work. For example, while the atmospheric data constraint in this inversion study does not cover the entire continent (Figure 6.2), an inter-comparison of inversion results with the forward models may still allow us to draw conclusions regarding strong sources or sinks that are visible from the atmospheric data in the well-constrained areas that are not apparent in the forward model estimates.

Results are shown in this section from both the FF-only and NARR inversions. The FF-only inversion provides a completely independent comparison to the forward models, while the NARR inversion may contain some variables also used as input into the forward models. However, given that the NARR variables are selected and their relationships to flux are inferred using the atmospheric data, the NARR inversion can also provide some insight into the forward model spread. In order to isolate the biospheric portion of the inferred flux, the fossil fuel inventory dataset is subtracted from the *a posteriori* flux estimates from the FF-only inversion, whereas the influence of the fossil fuels was already pre-subtracted from the measurements in the NARR inversion (Section 6.2.3.3).

6.4.1 Seasonal grid-scale spatial patterns

Figure 6.4 shows the three-monthly average fluxes from the FF-only and NARR inversions, as compared to the average NEE across the forward models included in the NACP Regional Interim Synthesis (Huntzinger et al., in prep.). The Root Mean Squared Difference (RMSD) and spatial correlation values between 3-monthly average fluxes

between the inversions and the forward model mean are shown for each season in Figure 6.4. The forward model mean, which masks the large spread in individual model results, was chosen for comparison, as this mean can be thought to represent the current "best guess" of the forward modeling community as a whole. This mean, which reduces the influence of outliers while still reflecting consistent patterns across models, also shows the lowest RMSD with the inversion at the monthly grid-scale (when compared to the RMSD with individual bottom-up models), and the second highest correlation (see Section 6.4.3). By comparing inversion results to the forward model mean, we help to identify features in the inversion that are more or less robust as compared to a set of process-based estimates. We also use results we believe to be robust from the inversion to hint at missing or misrepresented processes in the forward models.

As noted previously in Section 6.3.1, the atmospheric data constraint in 2004 does not cover the entire continent (Figure 6.2), and it changes from month to month. The grid-scale uncertainties, not shown here, reflect the available data, with higher uncertainties in areas weakly or not constrained at all by the atmospheric measurements, as shown in Mueller et al. (2008). As also discussed in Chapter 3, the estimated fluxes revert to the model of the trend ($\mathbf{X}\widehat{\boldsymbol{\beta}}$) in under-constrained areas. Therefore, in the FF-only inversion, the trend provides little information about the biospheric fluxes in under-constrained areas, where these biospheric fluxes revert to zero. For the NARR inversion, the trend is defined by the NARR auxiliary variables combined with their inferred drift coefficients ($\widehat{\boldsymbol{\beta}}$). To help avoid interpretation of fluxes exclusively defined by the trend in both inversions, fluxes from areas with almost no sensitivity to the measurements, i.e. Central America and Greenland, are not displayed in Figure 6.

With the FF-only inversion estimates, spatial patterns are relatively diffuse and unrealistically smooth as compared to the NARR inversion and the forward model mean (Figure 6.4). However, a comparison of these smooth spatial patterns across seasons can still help to shed light on the different model estimates. At the most basic level, the

seasonal cycle of sources and sinks is consistent between the FF-only inversion fluxes and the forward model mean, i.e. strong sinks during the growing season and sources during the dormant season. Also, the spatial patterns and magnitude of fluxes are most consistent between the inversion and the forward model mean during the height of the growing season from June to August when the flux signal is the strongest and most variable.

From December to February, the FF-only inversion results are particularly smooth, given that this season has the longest inferred flux correlation lengths (Figure 6.3a). The isotropic spatial covariance structure in $\mathbf{Q}$ (equal correlation lengths in all directions) may also be unrealistically extrapolating sources to the northern and southern portions of the domain, as compared to the forward model mean. The inclusion of NARR variables during this season is seen to have no impact on improving the correspondence with the forward model mean, perhaps because most of the selected NARR variables primarily help to explain fluxes during the growing season. Also, errors in appropriately separating the biospheric and anthropogenic signals in the total CO_2 flux may be most apparent during the winter, complicating the correlation with NARR variables.

The biospheric flux estimates from the FF-only inversion are mostly neutral from March to May, although the sources in the boreal North are somewhat inconsistent with the forward model mean, pointing to differences in the magnitude of respiration fluxes at the onset of the growing season in the bottom-up models. During this season, the inclusion of the NARR variables has a particularly strong impact on flux estimates with more clearly identifiable flux patterns across the continent relative to the FF-only inversion. With NARR datasets, realistic patterns are now seen in the under-constrained areas, e.g. sinks in the coastal Pacific forests and in the south-eastern continental United States. From March to May, the correlation coefficient with the forward model mean goes up from 0.34 to 0.67 when including NARR variables, the highest increase across the different seasons.

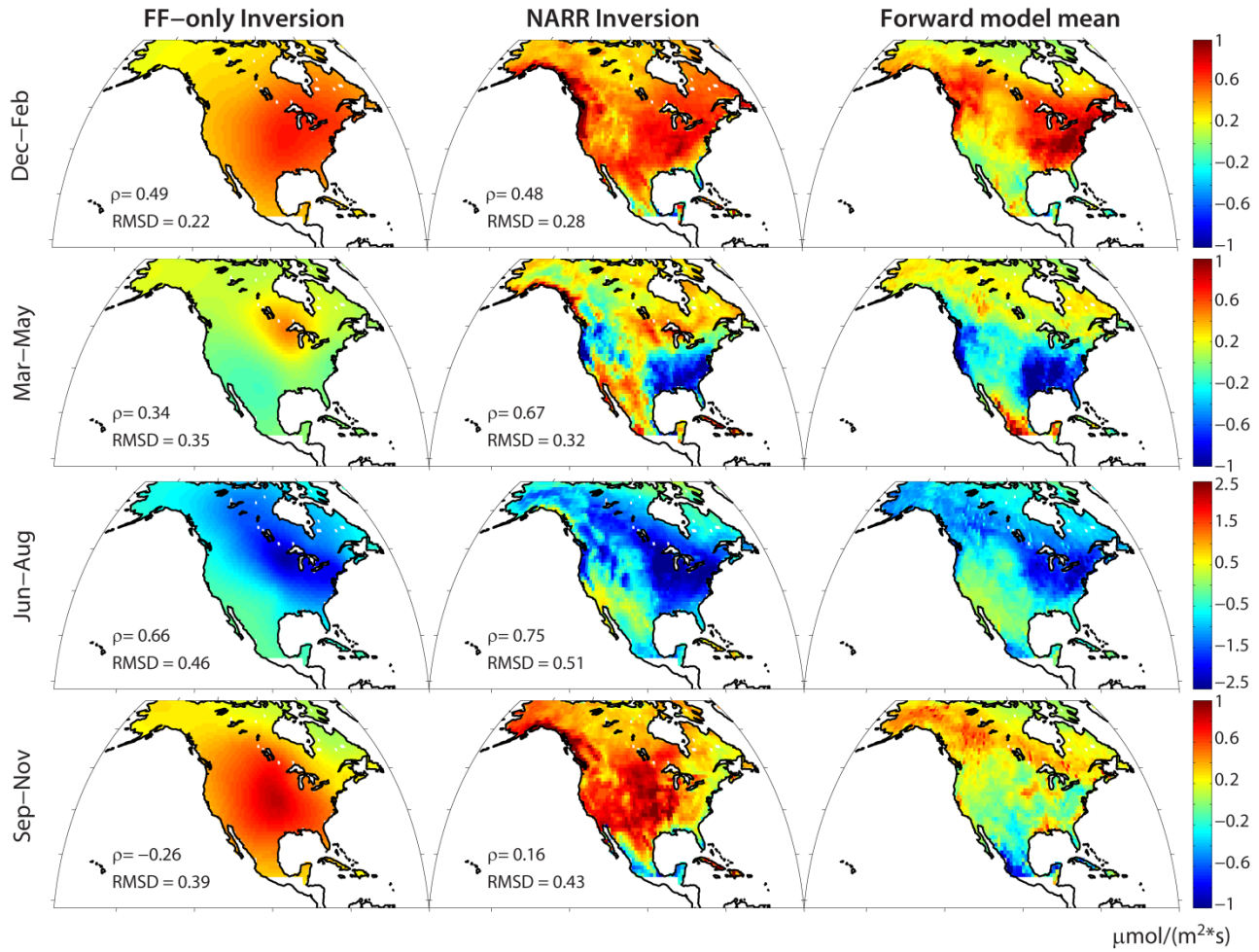

Figure 6.4: Three-monthly average grid-scale biospheric fluxes from the FF-only and NARR inversions with GlobalView boundary conditions, as compared to the forward model mean. The fossil fuel inventory is subtracted a posteriori from the estimated total CO_2 flux to isolate the biospheric contribution. Also shown are the spatial correlation (ρ) and Root Mean Squared Difference (RMSD) between the 3-monthly average inversion fluxes and the forward model mean. Please note the different scales for each season.

During the height of the growing season from June to August, the strong uptake in the agricultural Midwest and relatively neutral fluxes in the desert southwest from the FF-only inversion are consistent with the forward model mean, while the weaker sinks in Alaska during this season most likely point to a lack of atmospheric data constraint, with large data gaps at the BRW tower throughout the year. In the FF-only inversion results, stronger sinks are seen in the far boreal north and tundra as compared to the forward model mean, although the isotropic spatial correlation structure may again be affecting this result to some extent. After including NARR variables, the correlation with the forward model mean goes up from 0.66 to 0.75, the highest correlation between the inversion and the forward model mean across the four seasons. Also, the inversion now shows a sharper transition from carbon sinks to neutral fluxes across the continental divide, and weaker sinks in the tundra, helping to correct potential biases associated with the correlation structure in $\mathbf{Q}$.

From September to November, the FF-only inversion shows anti-correlations with the spatial patterns in the forward model mean. The anti-correlations in this season appear to be specifically driven by large sources in October in the center of the continent from the inversion, which are not seen in the forward model mean, or in the great majority of the individual forward model results. An analysis of the underlying concentration data shows a strong build-up of CO_2 relative to background air at all towers in the continental United States for this month, particularly at the LEF and WKT measurement towers, implying that these sources in the center of the continent are not an artifact of the particular inversion setup used here. Given that the central United States is a heavily agricultural area, this relatively robust feature of the inversion results may point to limitations in appropriately modeling agricultural CO_2 fluxes in the forward models (Lokupitiya et al., 2009; Corbin et al., 2010; Huntzinger et al., in prep.). The inclusion of NARR datasets improves the correlation with the forward model mean from -0.26 to 0.16 from September to November, although NARR inversion results are still the least correlated with the forward model mean during this season.

Overall, this inter-comparison of grid-scale spatial patterns across models and different seasons shows that the introduction of NARR variables into the linear model of the trend improves the ability of the inversion to recover realistic spatial patterns, particularly during the growing season, as shown by an improved correspondence with the forward model mean. Inversion results with and without NARR datasets also point to missing sources and/ or misrepresented processes in the forward models from September to November.

6.4.2 Magnitude and timing of biome-scale seasonal cycle

Figure 6.5 shows the monthly seasonal cycle for flux estimates aggregated to the seven biomes (Olson, 2001) shown in Figure 5.8, where these biomes define spatially continuous regions with similar land-cover and climatic characteristics. Figure 6.5 also includes the aggregated seasonal cycle for the full continent excluding Central America and Greenland, given that, in these areas, there is little sensitivity to the atmospheric measurements for the inversions, and many forward models do not have estimates.

The uncertainties shown in Figure 6.5, although inferred using an inversion run with a 4-day average diurnal cycle temporal flux resolution as discussed in Section 6.2.3.2, generally reflect the atmospheric data constraint, with wider confidence intervals in the under-constrained biomes. The narrower confidence intervals for North America indicate that the seasonal cycle at the continental scale is better-known than at finer spatial scales. While pseudo-data testing revealed that the confidence intervals for all regions shown here may be too narrow, they also confirmed the general conclusion that the aggregated seasonal cycle from the inversion is mostly robust in the well-constrained areas, i.e. the Eastern Temperate and Boreal Forests, the Temperate Grass, Savannah & Shrub (or the agricultural areas), and also for the continent. Pseudo-data tests did show that the inversion infers a later start to the growing season as compared to the “true” solution, but that the inclusion of auxiliary variables helps to correct this bias.

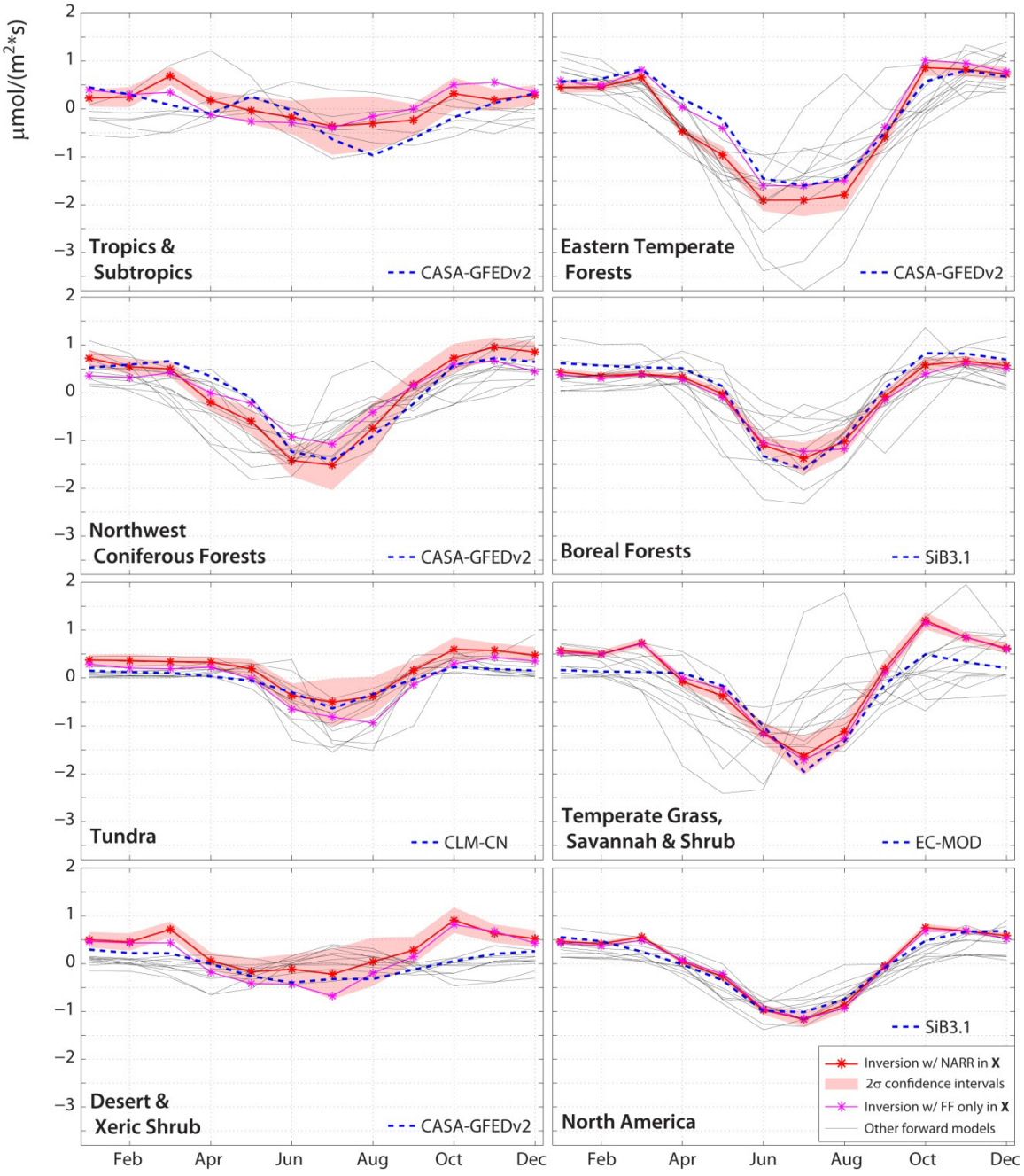

Figure 6.5: Seasonal cycle of monthly fluxes aggregated to seven eco-regions (shown in Figure 5.8) and the continent. Inversion fluxes are shown for the FF-only and NARR inversions with GlobalView boundary conditions. The $2\sigma_{\hat{s}}$ confidence intervals are from the NARR inversion. Results are compared to all forward models with at least 85% area coverage for the given biome or domain, and a specific model having the closest agreement with the inversions at this aggregated spatial and temporal scale is also highlighted for each region.

A comparison of inversion results in Figure 6.5 from the FF-only and NARR inversions shows that the inclusion of NARR auxiliary variables has little impact on flux estimates at this aggregated scale, consistent with results from Chapter 4. This shows that these environmental datasets do not bias flux estimates at aggregated spatial and temporal scales, where estimates are instead primarily constrained by the atmospheric data. This result is particularly true in well-constrained biomes, such as the Boreal Forest and the Temperate Grass, Savannah & Shrub, as seen in Figure 6.5. In the biomes with a weaker data constraint, e.g. the Tundra and Desert & Xeric Shrub, the NARR auxiliary variables have more of an impact, slightly nudging the inversion results towards the middle of the forward model spread, at least in some months of the year. The NARR auxiliary variables also lead to an earlier start to the growing season and a stronger magnitude of peak uptake in the Eastern Temperate Forests. Parts of this biome are well-constrained (see Figure 6.2), but this biome also includes the under-constrained southeastern United States forests, which are known to be very productive (Baker et al., 2010; Crevoisier et al., 2010).

Flux estimates from both inversions are also compared in Figure 6.5 to the individual forward models. In addition, one specific model is highlighted for each biome, where this model has the lowest combined rank in terms of RMSD and correlation with the inversions at this scale. Given that the inversion results are relatively robust at this aggregated scale in well-constrained areas, this inter-comparison helps to provide insight into the forward model spread. By highlighting an individual model, we also provide hints as to how a specific bottom-up model would need to be corrected in future work in order to be more consistent with the atmospheric data constraint.

With the exception of under-constrained biomes such as the Tropics and the Desert and Xeric Shrub, where unrealistic seasonality is most likely a result of extrapolation of inversion-derived fluxes from the better-constrained areas, Figure 6.5 shows that the inversions generally recover a seasonal cycle within the spread of the

forward models. That said, the forward model spread itself can be quite large, particularly in the Temperate Grass, Savannah and Shrub, i.e. the agricultural areas of the continent, and the Eastern Temperate and Boreal Forests, where the magnitude of peak uptake differs strongly among the models.

CASA-GFEDv2 shows relatively close agreement with the FF-only inversion in the Eastern Temperate Forest, especially in terms of the timing and magnitude of fluxes in the early growing season, although the NARR inversion shows stronger uptake from April to August. If the NARR inversion is more realistic in this area, particularly in terms of extrapolating the flux signal to the productive southeastern forests in the U.S., this may imply that CASA-GFEDv2 shows a late start to the growing season, as discussed in Randerson et al. (2009). The inversions also show particularly strong agreement with CASA-GFEDv2 in the Northwest Coniferous Forests in terms of the timing of the peak uptake, in distinct contrast to the majority of the forward models which shift the peak uptake one or two months earlier. In the Boreal Forests, SiB3.1 agrees very closely with the inversions, although with a slightly stronger seasonal cycle. Given that this is a relatively well-constrained biome, this result helps to give increased confidence in the estimates from SiB3.1 relative to other forward model estimates in this region.

The seasonal cycle in EC-MOD shows strong correspondence with the inversion from April through August in the Temperate Grass, Savannah & Shrub, or the agricultural regions, in terms of both the timing and magnitude of uptake. However, EC-MOD shows weaker sources than the inversion throughout the dormant season, particularly in March and October where the inversion shows sharp peaks in CO_2 release to the atmosphere. In fact, most of the forward models show weaker sources in March and October relative to the inversions, with the exceptions of CASA-GFEDv2, which appears to match the timing of the seasonal cycle fairly well from January to March, and DLEM, which shows strong sources in October, although with the peak source shifted one month later to November. The generally larger sources from the inversion throughout the dormant season in this biome may also point to errors in appropriately

separating the biospheric and fossil fuel signals, which would be most apparent when the fossil fuel emissions are dominating the total CO_2 flux.

At the continental scale, the spread in the forward model estimates is narrower, and the inversion agrees relatively closely with most models, with the exception again of the sources in March and October which are still visible at this highly aggregated spatial scale. Among the models, SiB3.1 has the closest agreement with the inversions for the continental seasonal cycle.

Overall, this comparison of the aggregated biome-scale seasonal cycle shows how the inversion can possibly help to give insight into the forward model spread in the well-constrained areas of the continent, as well as point to potential improvements in individual models. While pseudo-data testing showed that inversion results were mostly robust at this aggregated scale, implying that the inversion could be used as a validation tool for the forward models, systematic transport errors and/ or other problems with the inversion setup in a real-data environment could reduce the value of this inter-comparison.

Despite the caveats, a few interesting results emerge from this inter-comparison. For example, the forward models in general estimate weaker sources than the inversion in March and October in the central areas of the continental United States (e.g. Temperate Grass, Savannah and Shrub). This may again point to limited ability in the forward models to appropriately capture processes associated with planting and harvesting schedules in agricultural areas. In terms of individual models, CASA-GFED has similar timing as the inversions in the Northwest Coniferous and the Eastern Temperate Forests, although with a late start to the growing season in this latter biome. SiB3.1 appears to do an excellent job in capturing the seasonal cycle in the Boreal Forests, while EC-MOD matches the inversion closely in the Temperate, Grass, Savannah & Shrub during the growing season, but misses the sources evident in the atmospheric data during the dormant season.

6.4.3 Grid-scale statistical comparisons between inversion and other models

In Figure 6.6, we compare the grid-scale FF-only inversion results with individual forward models using statistical diagnostics. Biospheric flux estimates are compared for the entire continent using the grid-scale RMSD and correlation between the inversion and the forward models at both the (a) monthly and (b) annual timescales. The correlation at the monthly timescale is in both space and time, whereas the annual correlation is only in space. In addition, results from three inversions with flux estimates over North America are included for comparison, those from Schuh et al. (2010), CarbonTracker v. 2009 (Peters et al., 2007), and Butler et al. (2010) with two underlying sets of prior flux estimates. It should be noted that the grid-scale spatial patterns from the Peters et al. (2007) and Butler et al. (2010) studies are defined by their bottom-up prior flux estimates, with the inversion only adjusting these patterns at coarser spatial scales. The FF-only inversion was chosen here for this inter-comparison given its complete independence from the forward model estimates, although the results using the NARR inversion led to similar conclusions at the monthly time-scale. Differences at the annual time-scale between the two inversions will be discussed further below.

At the monthly timescale, one can see that the inversion has the strongest correlation and lowest RMSD with CASA-GFEDv2 among all the forward models, and even other inversions. While other forward models have a wide spread in terms of their correspondence with the inversion at this scale, one can also see that the forward model mean agrees closely, and in fact has a similar level of agreement with the inversion as CASA-GFEDv2. This result may help to validate the choice of CASA-GFEDv2 as an explicit prior flux estimate in the CarbonTracker data assimilation system (Peters et al., 2007). At the monthly timescale, no conclusions can be drawn as to the relative agreement of prognostic vs. diagnostic models, given that there are models that agree more or less strongly in each category.

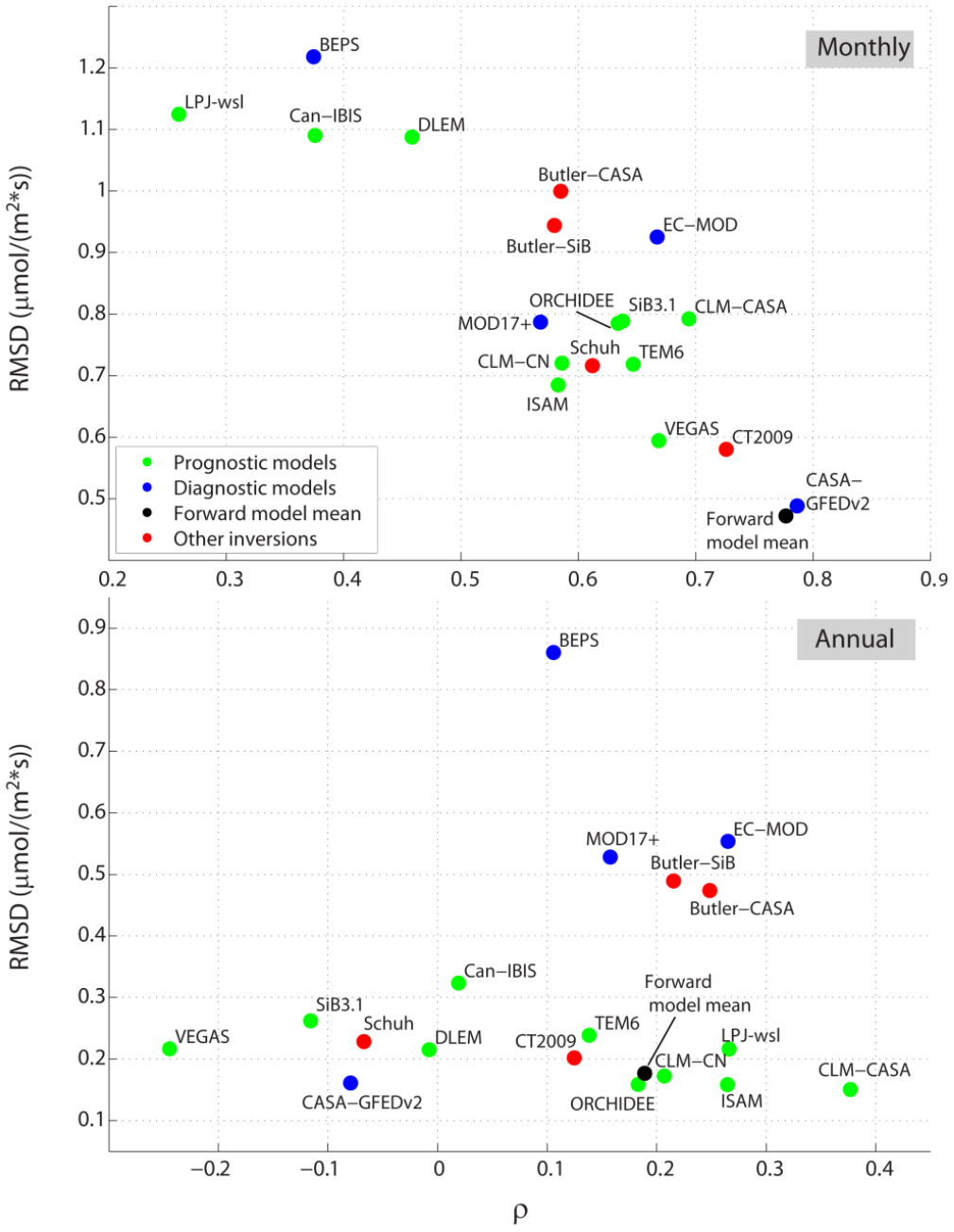

Figure 6.6: Root Mean Squared Difference (RMSD) and correlation between grid-scale biospheric fluxes for the continent from the geostatistical inversion and the forward models included in the NACP Regional Interim Synthesis, as well as other inversion studies as described in Section 6.4.3. Diagnostics were calculated at the monthly and annual timescales, where the correlation at the monthly timescale is in both space and time, and the annual correlation is just in space. Flux estimates from the FF-only inversion with GlobalView boundary conditions are used here for the inter-comparison.

At the annual timescale, results show lower RMSD values than at the monthly timescale, but also lower correlations between the inversion and the forward models. These lower correlations can be explained in a few ways. First, the annual-scale diagnostics are "spatial-only", whereas the monthly-scale diagnostics also capture the seasonal cycle which is relatively well-known. Second, the areas of the continent constrained by the atmospheric data in the inversion shift from month to month, and this, combined with systematic transport errors, most likely lead to larger biases in the inversion fluxes at the annual scale. Finally, it may be that the forward models themselves have less skill at capturing net annual sources and sinks, relative to their ability to accurately model monthly spatial patterns and the shape of the seasonal cycle.

The relative level of agreement between the individual forward models and the inversion also differs between the monthly and annual timescales. For example, SiB3.1, CASA-GFEDv2 and VEGAS2, which had relatively close agreement at the monthly time-scale, are anti-correlated with the inversion at the annual scale! To help identify whether this difference was due to the inclusion of the well-known seasonal cycle in the monthly timescale plot, the average monthly "spatial-only" correlation was calculated instead. These values were seen to be greater than zero for all the models, and greater than 0.3 specifically for SiB3.1, CASA-GFEDv2 and VEGAS2. Therefore, the agreement between the grid-scale spatial patterns in the inversion and these forward models is definitively poorer for the aggregated annual timescale, relative to the monthly results.

In addition, at the annual timescale, the prognostic models have lower RMSD, on average, than the diagnostic models. In contrast, the large RMSD values seen in BEPS, EC-MOD and MOD17+, three diagnostic models, remained high from the monthly to the annual timescales. Some of the prognostic models also improved their relative agreement with the inversion in terms of their correlation at the annual timescale. For example, CLM-CASA', a prognostic model, has the closest agreement with the inversion at the annual scale, while LPJ-wsl moves from being the least to the 2nd most correlated model from the monthly to the annual scales. Overall, these results suggest not only that model quality may differ across timescales (Stoy et al., 2009; Yadav et al., 2010;

Mueller et al., 2010), but that the reliance on remote-sensing datasets to help capture fluxes at short timescales in the diagnostic models may be at the expense of capturing longer-term processes that drive net annual sources and sinks.

Another interesting result from Figure 6.6 is that for all synthesis Bayesian inversions included in the inter-comparison, the geostatistical inversion sees a stronger correlation at the monthly timescale with the underlying biospheric model used as the explicit prior than the inversions themselves. (CarbonTracker used CASA-GFEDv2 as a prior, Schuh et al., 2010 used SiB3, and Butler et al., 2010 used both CASA-GFEDv2 and SiB3.) This result suggests that these previous inversion studies may be subject to aggregation errors due to their limited ability to correct fine-scale fluxes in both space and time around the towers. For example, CarbonTracker only has freedom to adjust the prior flux estimates at the weekly, eco-region scale over North America (Peters et al., 2007), whereas Butler et al. (2010) estimates flux corrections at the monthly timescale for 10 sub-regions on the continent. In contrast, at the annual timescale, these other inversion studies show relatively stronger agreement with the geostatistical inversion than their priors, suggesting that flux estimates at this scale are more driven by the atmospheric data constraint than the choice of bottom-up prior flux estimates.

The inclusion of NARR datasets into the inversion moderately improves the agreement with the forward models at the monthly timescale, without changing the relative position of the models (results not shown). However, including the NARR datasets into the inversion has the most impact at the annual scale. Despite a slightly higher RMSD, 13 of 15 forward models and all inversions show stronger annual-scale spatial correlation with the inversion after the inclusion of NARR auxiliary variables. For example, the correlation with CLM-CASA' goes up from 0.38 to 0.44 for the continent, with an improvement of 0.27 to 0.58 specifically in the high sensitivity areas shown in Figure 6.2. EC-MOD also becomes the most correlated model with the inversion at the annual scale (from 0.27 to 0.55) after including NARR data, although still with relatively high RMSD. The agreement with CASA-GFEDv2 goes up from -0.08 to 0.14, although it still remains the 5th least correlated model with the inversion at the annual scale.

Overall, it is difficult to know if the higher level of agreement between the NARR inversion and the forward models at both the monthly and annual timescales is mostly due to the fact that some of these same variables were also used as input into the forward models. If this is the case, then all models may be subject to correlated errors associated with biases in the input data. However, the NARR variables are incorporated into the inversion in a manner consistent with the atmospheric data constraint, providing additional information especially important in this data-limited year. Therefore, it seems likely that these variables are helping to improve flux estimates from the inversion to some extent, and that the closer agreement between the inversion with NARR variables and the forward models may be taken as evidence that both sets of models are starting to converge on a "true" solution.

6.5 Results: Annual carbon budget

At the annual time-scale, we present results at three different spatial scales. First, we examine the grid-scale spatial patterns as compared to two specific forward models. Next, we present spatially aggregated annual budgets for the continent, as well as for the high sensitivity areas shown in Figure 6.2, where we expect estimates to be mostly constrained by the atmospheric measurements year-round. Also, while inversion results using the GlobalView boundary conditions were exclusively presented in Section 6.4 for the inter-comparison of spatial patterns and the seasonal cycle of flux estimates, we present inversion results at the annual timescale using both the GlobalView and CarbonTracker boundary conditions.

6.5.1 Annual grid-scale sources and sinks

Figure 6.7 shows the annual grid-scale sources and sinks from the inversions with both sets of boundary conditions, as compared to EC-MOD and CLM-CASA'. Only NARR inversion results are shown given their greater correspondence with the forward models at this scale, as discussed in Section 6.4.3. EC-MOD and CLM-CASA' are also chosen for

comparison given their high correlations with the NARR inversion results at the annual grid-scale.

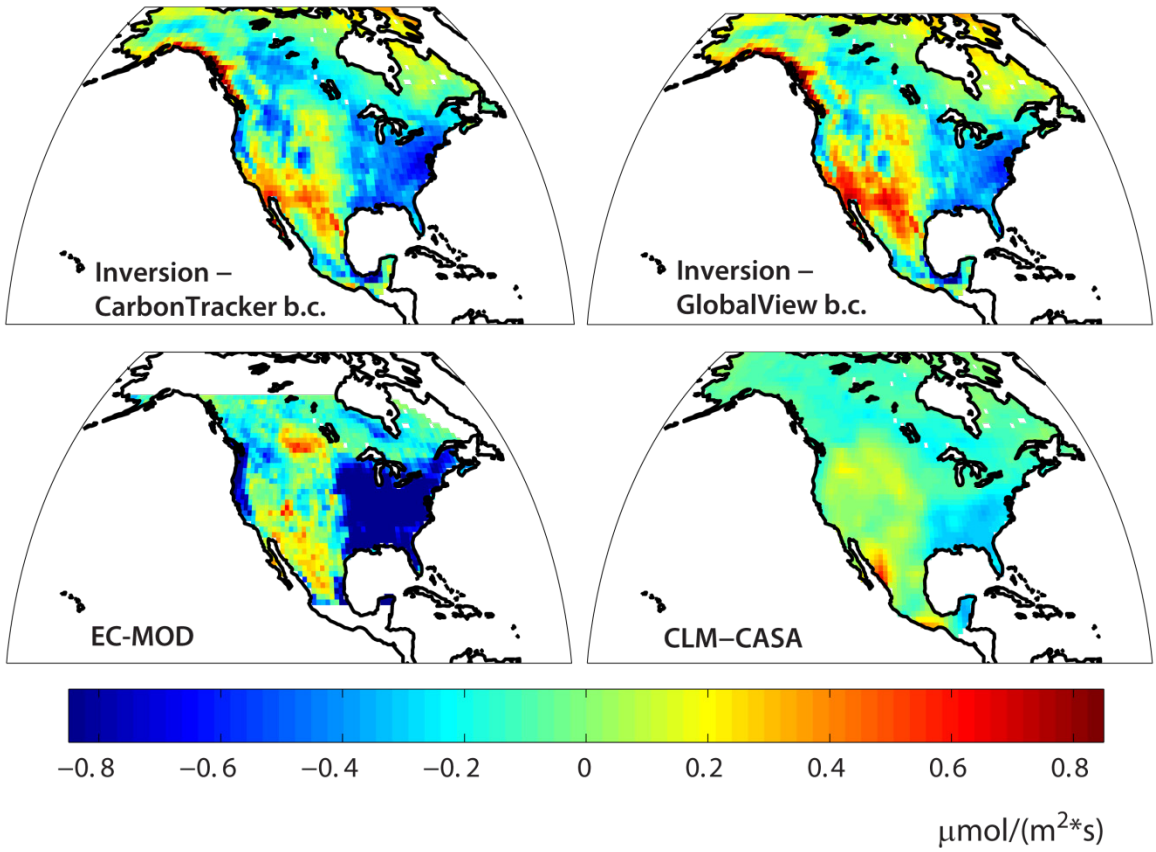

Figure 6.7: Grid-scale annual biospheric sources & sinks from the NARR inversion, using two sets of boundary conditions. Inversion fluxes are compared to annual flux estimates from EC-MOD and CLM-CASA'.

In Figure 6.7, the boundary conditions are shown to mostly impact the magnitude of flux estimates at the annual time-scale, rather than their spatial patterns. However, these magnitude differences can be relatively large. For example, the average annual grid-scale CO_2 flux for the Eastern Temperate Forests is -0.3 µmol/(m^2*s) using the GlobalView boundary conditions, but -0.4 µmol/(m^2*s) using the CarbonTracker dataset, a 33% increase. Also, the large sources in the desert Southwest with the

GlobalView boundary conditions are smaller in magnitude than results using the CarbonTracker dataset. An analysis of monthly grid-scale flux estimates from both inversions showed that most of the magnitude differences at the annual time-scale were due to stronger net sinks in the inversion using CarbonTracker boundary conditions from March through August. Magnitude differences in flux estimates were still evident, but smaller outside of the growing season. However, the consistent offset in the same direction between the two sets of boundary conditions at all towers throughout the year ensures that monthly differences in flux magnitudes become additive across time, rather than canceling each other out.

In terms of the net annual spatial patterns, inversions with both sets of boundary conditions show net uptake in the eastern continental United States and the boreal forests, as well as in southern Mexico and the Pacific coastal forests, although these latter regions are highly under-constrained by the atmospheric measurements. The locations of net uptake from the inversion are in relatively close agreement with the estimates of Crevoisier et al. (2010) from 2004 to 2006, who used an independent carbon budgeting method for North America based on vertical profiles of CO_2 concentrations over the continent, as well as bottom-up inventory estimates from the State of the Carbon Cycle Report (CCSP, 2007). The inversion also shows net sources in the desert Southwest, which is consistent in terms of sign with annual fluxes in most of the forward models, although the magnitudes from the inversion are much higher. This may point to errors in the forward models in this region, underestimation of fossil fuel emissions in the inventory, errors in the boundary conditions (particularly GlobalView in this case), or problems with the inversion setup, e.g. systematic transport errors at WKT or ARM. Given the limited vegetation in this area, Crevoisier et al. (2010) hypothesized that agricultural consumption in the pasture-lands in these areas could be contributing to neutral fluxes or net sources.

Figure 6.7 also shows that the spatial patterns between the inversions and EC-MOD have striking similarity, perhaps because EC-MOD uses a similar approach to the inversion in terms of extrapolating spatial patterns of flux with MODIS datasets (Xiao et

al., 2008). Xiao et al. (2008) also chose these datasets in a statistically rigorous manner, although using eddy-covariance flux tower measurements to train the model, rather than atmospheric measurements in an inversion framework as implemented here. Figure 6.7 also shows that, although EC-MOD and the inversions show similar spatial patterns, the magnitude of flux differs, i.e. EC-MOD shows much stronger sinks in the eastern half of the continental United States than the inversions, and weaker sources in the western half, with some consistency in the existence of sources in the desert Southwest. The annual-scale flux estimates from CLM-CASA' also show similar spatial patterns compared to the inversion, e.g. sinks in the central and southeastern United States, faint sinks in the boreal forests, and some sources in northwest Mexico, but in general the net annual fluxes are much smaller in magnitude than those from either the inversions or EC-MOD.

Overall this comparison of the net annual spatial patterns of CO_2 flux suggests some convergence in the location of net sources and sinks over the continent among models for this year. However, the divergence in flux magnitude across models still makes it difficult to provide carbon budgets with reasonable uncertainties at small regional scales, e.g. states or provinces, from either top-down or bottom-up approaches at the current time. A number of additional tests exploring the sensitivity of the geostatistical inversion presented here to data choices, flux temporal resolution, and temporal covariance assumptions also showed a convergence in the pattern of net sources and sinks at the annual scale, but strong differences in the magnitude of flux across inversion setups.

6.5.2 Spatially-aggregated annual budgets

Figure 6.8 shows annual flux estimates spatially aggregated to both the high sensitivity areas (6.10a) and the whole continent (6.10b), as compared to the suite of forward models and other inversions. Please note that some of the forward models and inversions included in the grid-scale and seasonal cycle inter-comparisons were excluded here, given that their flux domain covered less than 95% of either the high

sensitivity areas or the continent. Figure 6.8 shows the total CO_2 flux with a line for the contribution from the fossil fuels, as estimated by the combined fossil fuel inventory dataset used in this study. While fossil fuel emissions are generally considered to be well-known, especially at the annual scale, there still remains some uncertainty associated with the magnitude of these estimates (e.g. Francey et al., 2010). Therefore, we show total CO_2 flux, as this is what is actually seen by the atmospheric measurements, and is the most robust result from the inversion. Also, pseudo-data tests revealed that the annual totals in the high sensitivity areas are more reliably estimated than those for the whole continent; despite this, continental-scale estimates are still shown due to the larger scientific interest in understanding the total North American carbon budget.

The most striking result seen in Figure 6.8 is the difference in the annual totals associated with the two different sets of boundary conditions. While the high sensitivity areas show a significant net sink regardless of boundary conditions, not surprising given that these are mostly crop or forested areas, the GlobalView boundary conditions completely erase the North American biospheric carbon sink at the continental scale within the 95% confidence intervals. In addition, the geostatistical inversion results using CarbonTracker boundary conditions give an almost identical North American carbon sink to CarbonTracker (Peters et al., 2007) itself. These results suggest that the annual carbon budget for North America from regional inversions is strongly influenced by the continental boundary conditions, which in turn points to the need for more research into the most robust set of values to use before these models could potentially be used as a carbon budgeting tool. Even in global inversions, fluxes and transport in other parts of the world essentially provide a boundary condition to the domain of interest, and errors outside the domain can therefore introduce bias into the total budget in the specified region.

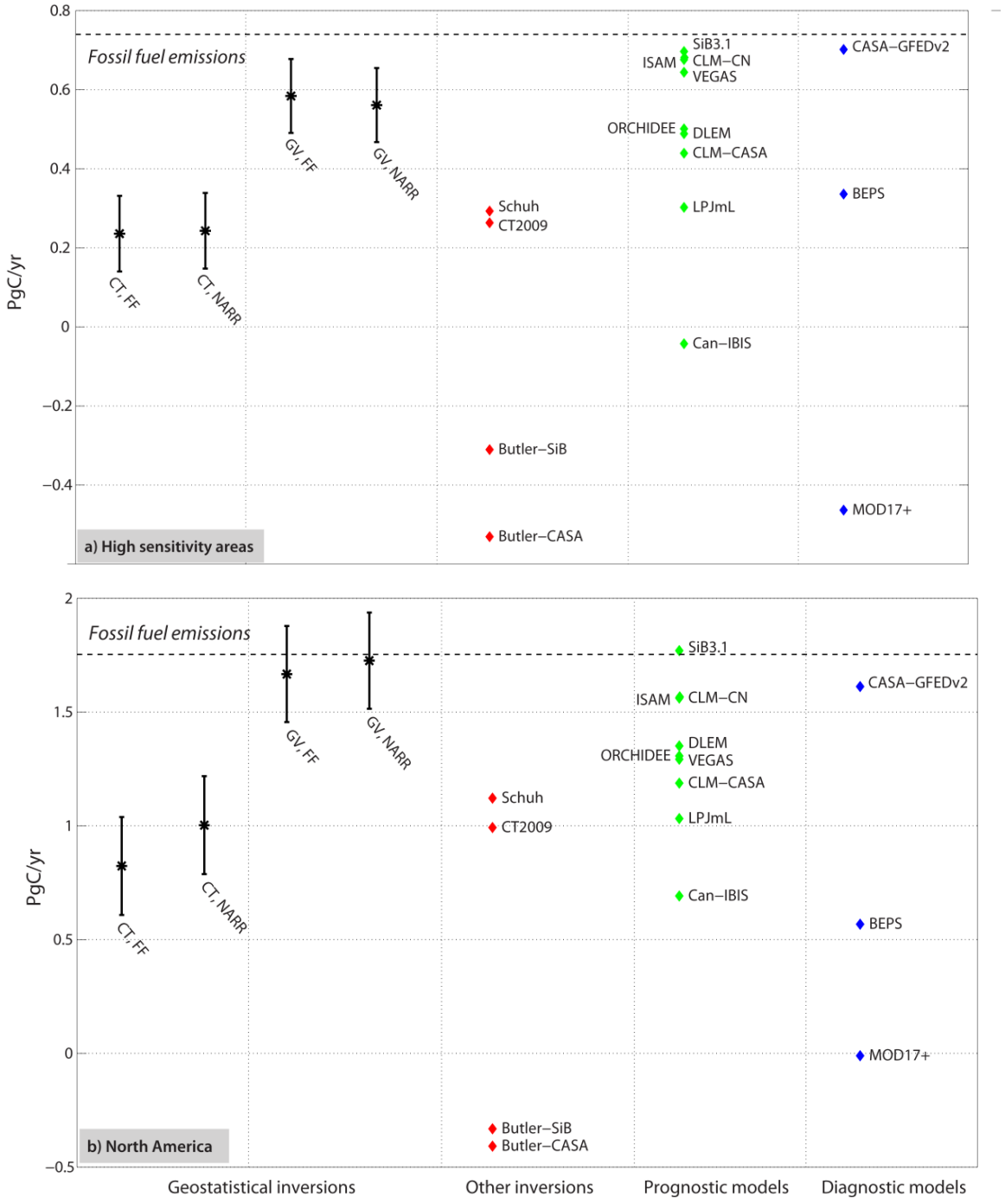

Figure 6.8: Annual total CO_2 source and sink estimates spatially aggregated to a) the high sensitivity area shown in Figure 6.2, and b) the continent. Results from the FF-only and NARR inversions using two sets of boundary conditions are compared to results from forward models and other inversion studies. The dotted line indicates the CO_2 flux from North American fossil fuel emissions, as estimated with the combined inventory dataset used in this study (Gurney et al., 2009; Oda and Maksyutov, 2010).

It should also be noted that in addition to the impact of the boundary conditions, the inferred biospheric carbon sink for the continent may also be biased by errors in the magnitude or spatial patterns of the fossil fuel inventory dataset, and by systematic transport and other errors in the inversion. For example, a sensitivity test excluding the 10am and 7pm data during the growing season increased the net sink in both the high sensitivity areas and the continent by 0.2 PgC/ yr. In contrast, results here are shown to be relatively insensitive to the components in the model of the trend for the geostatistical inversions. Therefore, these variables are useful for down-scaling the flux signal in the atmospheric data to finer spatial scales, for extrapolating to under-constrained areas, and for inference regarding flux drivers, but they do not bias flux estimates at larger spatial and temporal scales, as also seen in Chapter 4 (i.e. Gourdji et al., 2008).

When comparing the continental budget from the inversion with the GlobalView boundary conditions to the forward models, the inversion is seen to have close correspondence with the forward models that have more of a neutral biosphere (i.e. CASA-GFEDv2, SiB3.1, CLM-CN and ISAM). If the GlobalView boundary conditions are assumed to be more empirical, and therefore more "correct", as compared to those from CarbonTracker, the results shown here imply that the atmospheric constraint does not necessarily imply stronger net sinks than forward models, as seen in other bottom-up/ top-down inter-comparison studies over North America (e.g. http://nacp.ornl.gov/mast-dc/int_synth_contreg.shtml). However, in the high sensitivity areas, where pseudo-data testing showed more reliable annual budgets, the inversion results with GlobalView boundary conditions were closest to the forward models with intermediate sink magnitudes, e.g. Orchidee, DLEM and CLM-CASA'. If this result were transferable to the entire continent, this would imply a continental sink only slightly weaker than the results from CarbonTracker or the Schuh et al. (2010) study, perhaps around 0.5 PgC/yr, in excellent agreement with the results from Crevoisier et al. (2010).

The stronger net sinks seen in other North American inversions in Figure 6.8, particularly those of Butler et al. (2010), may potentially be biased by their boundary conditions and/ or fluxes in the rest of the world, the oversampling of relatively productive areas and the day-time photosynthesis signal within the North American continent, and aggregation errors that spread this sink bias out over large areas. Yet, without further research into the correct set of boundary conditions to use and the inclusion of additional measurement data from the post-2004 expanded measurement network, it remains difficult to accurately estimate the North American continental carbon budget using the atmospheric data constraint from any inversion study.

6.6 Conclusions

The geostatistical inversion approach implemented here resolves fluxes at finer scales in both space and time than other published inversion studies estimating fluxes over the North American continent (Peters et al., 2007; Deng et al., 2007; Schuh et al., 2010; Butler et al., 2010). By resolving fluxes at finer scales, this inversion reduces spatial and temporal aggregation errors associated with using continental measurements sited in areas with high flux variability. Also, the geostatistical inversion eliminates the requirement for explicit prior flux estimates allowing for a more independent comparison to bottom-up model output.

One encouraging sign in regards to avoiding systematic transport model errors in the inversion, which have the potential to significantly bias flux estimates, is that the recovered $\widehat{\boldsymbol{\beta}}$ on the fossil fuel inventory dataset was almost exactly one. This implies that the inversion is able to independently identify the fossil fuel signal in the total CO_2 flux (assuming that the spatial patterns and magnitudes in the inventory are basically correct), which provides hope for future work using the atmospheric constraint to validate the regional emissions and individual sectors that make up these inventories.

Introducing NARR auxiliary variables into the inversion was seen to help extrapolate the flux signal from the atmospheric data in a realistic way, as compared to the forward model mean from the NACP Regional Interim Synthesis. However, given

the limited atmospheric data constraint noticeable in the spatial patterns of the FF-only inversion (Figure 6.4), the drift coefficients ($\widehat{\boldsymbol{\beta}}$) on these variables may be somewhat biased by the relatively productive regions of the continent sampled by the network. Future inversions using the expanded in-situ measurement network over the continent (Mueller et al., in prep.) may help to remedy this situation. Regardless, the selected variables and inferred drift coefficients on the NARR variables were very consistent with process-based understanding of the drivers of both photosynthesis and respiration, providing further validation of the inversion setup implemented in this study.

Despite the lack of direct validation data for biospheric CO_2 fluxes at regional scales, the top-down/ bottom-up inter-comparison implemented in this study helped to provide insights into the strengths and weaknesses of both the inversion and the forward models. One conclusion was that the inversion had the most consistent spatial patterns with the forward model mean in the growing season relative to the dormant season. This could be due to errors in the fossil fuel inventories, which are more evident in the dormant season when emissions dominate the total CO_2 flux, or the forward models themselves may have less skill outside of the growing season (e.g. Schwalm et al., 2010). This result may also be related to the fewer number of measurements in the inversion during the dormant season when days are shorter, and there are fewer hours in the day with well-mixed conditions. Finally, the fluxes during the growing season are stronger and more variable, and therefore perhaps easier to identify from the atmospheric signal.

Another conclusion that emerges from the top-down/ bottom-up inter-comparison in this study is that there are individual forward models that appear to agree well with the inversion in specific seasons, biomes or scales of comparison. However, there is no one model that appears to agree all the time with the inversions. For example, CASA-GFEDv2 has the closest agreement in terms of RMSD at the monthly timescale (including both spatial patterns and the seasonal cycle). However, the grid-scale spatial patterns of net annual fluxes in CASA-GFEDv2 were seen to be anti-correlated with the inversion results! Similarly, EC-MOD and CLM-CASA' have closer

correspondence with the inversion at the annual, relative to the monthly, timescale. While a lack of agreement may imply errors in either the inversion or the forward models, pseudo-data testing did reveal that inversion results at aggregated scales within the well-constrained areas were relatively robust. Therefore, the inversion was found to be useful for giving insight into the forward model spread, as demonstrated for the biome-scale seasonal cycle.

One result from the inversions that was robust across all sensitivity tests (varying the data inputs, flux resolution and temporal covariance assumptions) were the strong sources in the center of the continent in March and October. While the magnitude of these sources was somewhat sensitive to inversion setup, a peak in respiration sources in the same regions during these months was seen in very few of the forward models. This may confirm previous work suggesting that biospheric models in general need significant improvements in agricultural areas (Lokupitiya et al., 2009; Corbin et al., 2010).

Finally, accurate annual carbon budgets from either regional inversions or bottom-up models appear to remain an elusive goal at the current time. While there appears to be some convergence in the location of net uptake and release over the continent in the inversions, EC-MOD and CLM-CASA', the magnitudes of flux estimates were very different across models. Furthermore, the boundary conditions used as input into the inversions led to strikingly different annual carbon budgets for the continent, making it difficult to highlight agreement between the inversion and the forward models at this scale.

Overall, the North American geostatistical inversion for 2004 presented here shows promise for providing a robust inversion framework that can scale well with the recent expansion of the in-situ CO_2 monitoring network over North America. Results show that the flux signal sampled by the 2004 monitoring network is sufficient to accurately identify the fossil fuel emission signal, while also constraining biospheric flux estimates in a manner consistent with process-based understanding. Top-down/ bottom-up inversion studies like this one can also help to provide insight into the large

spread of bottom-up model estimates of regional CO_2 flux. Such comparisons can therefore help to identify needed improvements in individual model formulation, in hopes that modeled CO_2 flux estimates at regional scales from both the top-down and bottom-up approaches may at some point converge.

CHAPTER 7

Conclusions and future directions

This concluding chapter summarizes the objectives and findings of each of the three major components of the dissertation, acknowledges collaborations with fellow researchers, and also suggests directions for future work.

7.1 Contributions of dissertation

7.1.1 Global geostatistical inversion study

The first objective of this dissertation applied the geostatistical inverse modeling approach to a global inversion with real data, using the atmospheric data constraint along with auxiliary environmental variables from remote-sensing and socioeconomic inventory datasets to constrain flux estimates. Overall, results in this work, and in the companion paper (Mueller et al., 2008), showed that the geostatistical inversion approach is a viable alternative to synthesis Bayesian inversions for estimating continental-scale fluxes around the globe, while minimizing the number of process-based assumptions included in other published inversion studies. Also, by resolving fluxes at the relatively fine spatial resolution of the transport model, this approach reduced aggregation errors (Kaminski et al., 2001) associated with resolving fluxes for larger regions, while fixing spatial patterns within the region from a bottom-up model (e.g. Gurney et al., 2004; Baker et al., 2006). Geostatistical inversion results were shown to provide intermediate values between results from two other global inversion studies (Baker et al., 2006; Rödenbeck et al., 2003), providing some indication that the presented inversion was able to reduce biases associated with spatial aggregation errors and the use of explicit prior flux estimates in these other studies (Mueller et al., 2008).

Furthermore, incorporating auxiliary variables into the global inversion, which were defined at the grid-scale from remote-sensing and socioeconomic inventory datasets, helped to recover realistic grid-scale heterogeneity and extrapolate the flux signal in the atmospheric measurements to under-constrained areas. Also, flux estimates at the continental-scale were relatively consistent between two inversions with and without auxiliary variables, showing that flux estimates at large spatial scales from a geostatistical inversion are primarily constrained by the atmospheric measurements, especially in well-constrained areas. Therefore, this approach provides a means to take advantage of process-based information, contained in remote-sensing and inventory datasets, and to downscale and extrapolate flux patterns in a manner consistent with the atmospheric data constraint, while not introducing biases at larger spatial scales.

7.1.2 Regional synthetic data inversion study

The second component of this dissertation extended the geostatistical inversion framework for estimating sub-continental scale fluxes in a regional inversion over North America using continental, continuous measurements of atmospheric CO_2. Given potential biases associated with using fixed patterns to describe fine-scale spatial and temporal variability within the flux signal, this component developed an inversion setup that could estimate this variability directly without the use of bottom-up model output. Estimating the diurnal variability in the fluxes was shown to be critical in a pseudo-data environment for inferring unbiased post-aggregated monthly, ecoregion-scale fluxes. Also, this component of the dissertation showed that it was possible to reliably estimate covariance parameters for the inversion with the atmospheric measurements, further reducing reliance on assumptions contained within bottom-up models for parameterizing the inversion. Finally, this component investigated the impact of *a priori* temporal covariance assumptions on final flux estimates and uncertainties. While introducing minimal bias into the flux solution in the pseudo-data environment,

temporal covariance assumptions were shown to be critical for recovering appropriately wide confidence intervals at aggregated temporal scales.

7.1.3 Real data inversion for North America in 2004

The third component of this dissertation, the real-data North American inversion for 2004, first showed that reasonable flux estimates could be inferred for the full year using solely the atmospheric data constraint with the inversion setup developed in the second component. Furthermore, this component showed that by incorporating diurnally-varying auxiliary datasets into the inversion at the regional scale, a geostatistical inversion provides a viable means for directly estimating spatial patterns over the continent without relying on fixed fine-scale estimates from a biospheric model, which are likely to be incorrect in certain areas, seasons and times of the day given the large spread in individual forward model estimates. In particular, auxiliary variables from the North American Regional Reanalysis (NARR; Mesinger et al., 2006) appeared to do an excellent job in terms of extrapolating and down-scaling the flux signal over the continent, as seen by closer correspondence between inversion flux estimates and the mean of a collection of forward models from the North American Carbon Program Regional Interim Synthesis (Cook et al., 2009, Huntzinger et al., in prep). That being said, it is impossible to independently verify any of these individual model results, and the possible inclusion of NARR datasets as inputs into the forward models themselves reduced the independence of this comparison, potentially leading to correlated errors.

Overall, the comparison between the geostatistical inversions, with and without NARR variables, and the forward model mean showed some convergence in terms of the inferred seasonal spatial patterns, particularly during the growing season. During the dormant season, the lack of agreement between the inversions and the forward model mean could be due to lack of skill in the forward models outside of the growing season (e.g. Schwalm et al., 2010), errors in the fossil fuel inventories, or perhaps both.

In particular, the inversion results showed strong sources in the center of the continent in October at both the grid and aggregated biome-scale, which were seen in very few of the forward models.

Model inter-comparisons at the annual timescale showed that the magnitude of net annual sources and sinks remains difficult to pinpoint from either top-down or bottom-up approaches at the current time, a difficult task in general given the small magnitude of net annual fluxes as compared to the seasonal cycle of individual flux components. The boundary conditions in the inversion were also shown to have a large impact on the continental carbon budget, leading to very different conclusions in terms of the forward models that agree most closely with inversion results at this scale. Relying on the new empirically-based GlobalView data product for North American boundary conditions, the inversion returned a weaker continental sink for this year than that seen in previous published inversion studies.

The real-data inversion results for 2004 more generally showed the constraint of the 2004 continuous measurement network on North American CO_2 fluxes at various spatial and temporal scales. For example, a number of sensitivity tests indicated that the spatial patterns of flux could be resolved more clearly by adding more data throughout the day and night into the inversion, although the impact of systematic transport model errors could also be clearly seen in the resulting flux estimates. A careful use of *a priori* temporal covariance assumptions also showed promise in recovering more realistic grid-scale spatial patterns without the inclusion of NARR auxiliary variables. Another promising result was that the in-situ measurement network for this year within the geostatistical inversion framework could independently identify the signal associated with fossil fuel emissions in the total CO_2 flux, as seen by a recovered drift coefficient ($\widehat{\boldsymbol{\beta}}$) near one. This provides hope for validating fossil fuel emission inventories, or individual sectors, in future work.

7.1.4 Overall contributions

The three components of this dissertation, in sum, helped to develop the geostatistical atmospheric inversion framework for estimating CO_2 fluxes directly at fine spatiotemporal scales, while minimizing process-based assumptions inherent to synthesis Bayesian inversion approaches. Also, by eliminating the rigid constraints provided by mechanistic forward models used as explicit prior flux estimates in other inversion studies, this work provides a relatively independent comparison to such models for helping to ultimately close the gap between bottom-up and top-down understanding of CO_2 flux. Finally, the infrastructure developed as part of this dissertation, in terms of the transport model runs and code developed for data processing, running the inversion, RML and variable selection algorithms (described further below), as well as the theoretical inversion framework itself, will hopefully help to inform future work using in-situ and satellite-based CO_2 measurements to estimate regional-scale CO_2 fluxes.

7.2 Collaborations

The inversion work that forms the core of this dissertation was performed most closely in collaboration with Kim Mueller, my fellow Ph.D. student at the University of Michigan. First, the global inversion studies (Mueller et al., 2008; Gourdji et al., 2008) were published together as a two-part paper. Second, her feedback formed a valuable contribution towards helping to develop a regional inversion setup in the second and third components that could ultimately recover reasonable flux estimates. The second and third components of this dissertation are, in addition, helping to inform Kim's work on the impact of the expanding *in situ* measurement network on regional inversions in North America (Mueller et al., in prep). Arlyn Andrews at the National Oceanic and Atmospheric Administration (NOAA) also gave valuable feedback throughout on avoiding various pitfalls in regional inversions.

The WRF-STILT transport model runs for the second and third components were done in close collaboration with colleagues at NOAA and AER Corp. The WRF model

output was generated directly by Thomas Nehrkorn and John Henderson at AER, while Thomas, John and Janusz Eluskiewicz also generated a streamlined set of scripts for running WRF-STILT on the NASA supercomputers. Adam Hirsch at NOAA helped to develop the original scripts and run the transport model in the early stages of the project, while the final production runs of WRF-STILT were performed on the NASA supercomputers by Kim and me, with the additional assistance of Mike Trudeau and Gaby Petron at NOAA.

Finally, Vineet Yadav and Deborah Huntzinger, postdoctoral fellows at the University of Michigan, also contributed their time and efforts to the North American regional inversions. Vineet helped to make the estimation of 8 million fluxes for a single year computationally possible by optimizing the inversion, RML and variable selection code using linear algebra identities to reduce needed computations, and by implementing parallel and distributing computing techniques. In particular, Vineet, in collaboration with Kim, wrote the scripts for running real-data inversions on the NASA supercomputers in an efficient manner in the final months of the dissertation. Debbie provided a valuable link to the forward modeling community, particularly for the third component of the dissertation.

Finally, this dissertation had wider impacts beyond the specific work presented here. For example, the scripts for running WRF-STILT on the NASA supercomputers are still being used by other members of Dr. Anna Michalak's research group and colleagues at NOAA to investigate transport model errors (A. Andrews, M. Trudeau and G. Petron), as well as the potential placement of new measurement towers in North America for CO_2 monitoring (Yoichi Shiga). The WRF-STILT concentration footprints for 2004 and 2008 and WRF meteorological fields for 2004 through 2008 are now freely available for download (http://puorg.engin.umich.edu) to the wider community for other atmospheric applications over the North American continent. The code generated for the second and third components of this dissertation, written by Vineet and myself, is also well-documented and freely available for download from http://puorg.engin.umich.edu.

7.3 Future work

7.3.1 Direct extensions to current work

While the inversion work performed for this dissertation was mainly intended to help identify a reliable inversion setup that could be informative for future operational CO_2 inversions, there are a few direct extensions of the current work that would help to make it more directly applicable to the larger carbon cycle science community. For example, the introduction of temporal covariance assumptions into the real-data 2004 inversion was not computationally feasible for generating annual uncertainties using a 3-hourly flux temporal resolution. However, this *a priori* covariance was also shown to help fill in spatial patterns in the inferred flux estimates, without relying on auxiliary variables that reduce the independence of the comparison to the forward models. While these temporal covariance assumptions must be used with caution in times of the year with rapid changes, e.g. in early spring, they may also help to further provide insight into the forward model spread over North America. This topic warrants further investigation in order to improve flux estimates and recover reliable uncertainties that can help to strengthen conclusions from future top-down/ bottom-up inter-comparisons of CO_2 flux.

Given that 2004 was a very data-limited year, in comparison to the current time when there are more than 40 towers collecting continuous CO_2 concentration measurements over the North American continent, it would make sense to repeat many of the analyses for this dissertation using an inversion for a later year with more available measurement data. These analyses would help to confirm the conclusions of this dissertation, while also pointing to results that may have been specific to a data-limited setup. In fact, my colleague Kim Mueller has begun to investigate the impact of flux temporal resolution, data choices, and boundary conditions on geostatistical inversion results using data from 35 towers in 2008 (Mueller et al., in prep.). Preliminary results highlight the increased impact of temporal aggregation errors on inversion results using a denser measurement network, and also the value of adding

more measurement data in time into the inversion, in order to optimally take advantage of new measurement sites for constraining flux estimates at finer spatial scales.

Future work using the 2008 inversion setup includes the incorporation of NARR auxiliary variables to test the consistency of selected variables and inferred drift coefficients ($\widehat{\boldsymbol{\beta}}$) between 2004 and 2008, and also the incorporation of individual sectors from the Vulcan fossil fuel inventory database (Gurney et al., 2008) to potentially investigate the reliability of these inventories as seen by the atmospheric data constraint. Finally, the 2004 and 2008 inversions could easily feed into a longer multi-year inversion incorporating auxiliary variables, where these variables could help to provide inference regarding the drivers of inter-annual CO_2 flux variability over North America.

7.3.2 Larger community-wide future directions for regional carbon budgeting

Apart from specific extensions to the current work implemented in this dissertation, there appears to be a need for a more coordinated approach in the carbon cycle science community for producing reliable model output which can be used for regional carbon budgeting. A comparison of model output from both forward models and inversions currently shows a wide spread among and between the two types of models (Huntzinger et al., in prep., NACP Regional Interim Synthesis, http://nacp.ornl.gov/mast-dc/int_synth_contreg.shtml). While this dissertation hopefully provided some insight into an optimal inversion setup for taking advantage of surface flux information contained in atmospheric CO_2 concentration measurements, there is still a need for more in-depth analysis and controlled experiments involving different modeling groups. Such a coordinated approach would make it easier to identify which features of the different modeling approaches appear to be most robust, while reducing redundant efforts.

Also, in the case of inversions, there is a need to improve transport modeling capabilities in order to eliminate systematic biases that affect inversion results. For example, numerous sensitivity tests for the third component of this dissertation showed

that increasingly well-resolved spatial patterns could be recovered (even without NARR auxiliary variables) as more measurements throughout the day were added into the inversion. Improvements in transport model quality would make it feasible to incorporate more data, e.g. night-time measurements, into the inversion without the risk of biasing flux estimates. However, currently, there appears to be limited funding and support within the community for building and improving this basic infrastructure which are critical for future operational CO_2 inversions (A. Andrews and M. Trudeau, personal communication). Similarly, there is a need for well-documented and robust coding practices that can make it easier for research groups to easily share and understand one another's code.

In summary, a community-wide effort is needed to find a consensus method for optimally ingesting both in-situ and satellite measurements of atmospheric CO_2, as well as process-based information, into an inversion system that can provide operational regional carbon budgets around the globe and for specific domains like North America. Along with robust uncertainties associated with the resulting flux estimates, such a system can hopefully help to support and inform climate change mitigation policies by providing an understanding of the link between carbon management policies and resulting atmospheric CO_2 concentrations. For example, without a way to reliably confirm bottom-up estimates of fossil fuel emissions from the point of view of the atmosphere, it will be difficult to enforce future policy commitments for reducing these emissions (NRC, 2010). Also, the atmospheric constraint could be used as a means to test the effectiveness of land management strategies designed to increase carbon sequestration in forests and on agricultural lands. Finally, a better understanding of the drivers of regional-scale CO_2 exchange, which can mostly come through improving models as in this dissertation, should help to improve projections of the future response of the terrestrial biosphere and oceans to already-committed climate changes.

REFERENCES

Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P. P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin (2003), The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present), *J Hydrometeorol, 4*(6), 1147-1167.

Ahmadov, R., C. Gerbig, R. Kretschmer, S. Körner, C. Rödenbeck1, P. Bousquet, and M. Ramonet (2009), Comparing high resolution WRF-VPRM simulations and two global CO_2 transport models with coastal tower measurements of CO_2, *Biogeosciences*, 6, 807–817.

Akaike, H. (1974), A new look at statistical model identification, *IEEE Transactions on Automatic Control*, AC19, 6,716-723.

Alig, R.J. & B.J. Butler (2004), Projecting large-scale area changes in land use and land cover for terrestrial carbon analyses, *Environmental Management*, 33, 4, 443-456.

Andres, R. J., G. Marland, I. Fung, and E. Matthews (1996), A 1°x 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950 – 1990, *Global Biogeochem. Cycles*, 10(3), 419– 430, doi:10.1029/96GB01523.

Angert A., S. Biraud, C. Bonfils, W. Buermann, & I. Fung (2004), CO_2 seasonality indicates origin of post-Pinatubo sink, *Geophys. Res. Lett.*, 31,11,L11103.

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

Baker, I.T., A.S. Denning, & R. Stockli (2010), North American gross primary productivity: regional characterization and interannual variability, *Tellus B*, 62, 5, Sp.Iss. SI, 533-549.

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*.

Baker, J.M., T.E. Ochsner, R.T. Venterea, & T.J. Griffis (2007), Tillage and soil carbon sequestration – what do we really know? *Agriculture, Ecosystems and Environment*, 118, 1-5.

Bakwin, P.S., K.J. Davis, C. Yi, S.C. Wofsy, J.W. Munger, L. Haszpra & Z. Barcza (2004), Regional carbon dioxide fluxes from mixing ratio data. *Tellus,* 56B, 301-311.

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

Baldocchi, D., B.B. Hicks, & T.P. Meyers (1988), Measuring biosphere-atmosphere exchanges of biologically related gases with micrometeorological methods. *Ecology*, 69: 1331-1340.

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X. H., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K. T. P., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S. (2001), FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, *Bull. Am. Meteorol. Soc*., 82(11), 2415–2434, doi:10.1175/1520-0477.

Baldocchi, D. (2003), Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. *Global Change Biology*, 9, 479-492.

Baldocchi, D. (2008). Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Austr. J. Botany* 56, 1–26.

Bjorkman, M.P., E. Morgner, E.J. Cooper, B. Elberling, L. Klemedtsson, and R.G. Bjork (2010), Winter carbon dioxide effluxes from Arctic ecosystems: an overview and comparison of methodologies, *Glob. Biogeochem. Cyc.*, 24, GB3010.

Bonan, G. (2008), *Ecological Climatology*, Cambridge University Press, Cambridge, MA, 2nd ed..

Bondeau, A., P.C. Smith, S. Zaehle, S. Schaphoff, W. Lucht, W. Cramer, and D. Gerten (2007), Modeling the role of agriculture for the 20th century global terrestrial carbon balance, *Global Change Biology*, 13, 3, 679-706.

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), 1342 – 1346, doi:10.1126/science.290.5495.1342.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet & P. Monfray (1999), Inverse modeling of annual atmospheric CO_2 sources and sinks 1. Method and control inversion. *J. Geophys. Res.*, 104(D21), 26161-26178.

Brenkert, A. (1998), Carbon dioxide emission estimates from fossil-fuel burning, hydraulic cement production, and gas flaring for 1995 on a one degree grid cell basis, *Tech. Rep. NDP-058A,* Carbon Dioxide Information Analysis Center.

Bruhwiler, L.M.P., A.M. Michalak, D.F. Baker, and P.P. Tans (2005), An improved Kalman Smoother for atmospheric inversions, *Atm. Chem. Phys.*, 5, 2691-2702.

Buermann, W., B.R. Lintner, C.D. Koven, A. Angert, J.E. Pinzon, C.J. Tucker, and I.Y. Fung (2007), The changing carbon cycle at Mauna Loa Observatory, *Proc. Nat'l. Acad. Sci.*, 104, 11, 4249-4254.

Butler, M. P., K. J. Davis, A. S. Denning, and S. R. Kawa (2010), Using continental observations in global atmospheric inversions of CO_2: North American carbon sources and sinks. *Tellus B*, doi: 10.1111/j.1600-0889.2010.00501.x.

Canadell, J.G., P. Ciais, S. Dhakal, H. Dolman, P. Friedlingstein, K.R. Gurney, A. Held, R.B. Jackson, C. Le Quere, E.L. Malone, D.S. Ojima, A. Patwardhan, G.P. Peters, M.R. Raupach (2010), Interactions of the carbon cycle, human activity, and the climate system: a research portfolio, *Current Opinion in Environmental Sustainability*, 2(4): 201-311.

Canadell, J.G., C. Le Quere, C., M.R. Raupach, C.B. Field, E.T. Buitenhuis,P. Ciais, T.J. Conway, N.P. Gillett, R.A. Houghton, 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): 18866-18870.

Carouge, C., P. Bousquet, P. Peylin, P.J. Rayner, and P. Ciais (2010a), What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes - Part 1: Potential of the network, *Atm. Chem. Phys.*, 10, 6, 3107-3117.

Carouge, C., P. Peylin, P.J. Rayner, P. Bousquet, F. Chevallier, and P. Ciais (2010b), What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes - Part 2: Sensitivity of flux accuracy to inverse setup, *Atm. Chem. Phys.*, 10, 6, 3119-3129.

Climate Change Science Program (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.

Central Intelligence Agency, U. S. (2007), World Factbook (available at http://www.cia.gov/cia/publications/factbook/).

Chevallier, F., et al. (2010), CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *Journal of Geophysical Research-Atmospheres* 115, D21307, doi:10.1029/2010JD013887.

Chevallier, F., N. Viovy, M. Reichstein, & P. Ciais (2006), On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes. *Geophysical Research Letters*, 33, L13802, doi:10.1029/2006GL026496.

Ciais, P., J.G. Canadell, S. Luyssaert, F. Chevallier, A. Shvidenko, Z. Poussi, M. Jonas, P. Peylin, A.W. King, E.D. Schulze, S. Piao, C. Rödenbeck, W. Peters, F.M. Breon (2010), Can we reconcile atmospheric estimates of the Northern terrestrial carbon sink with land-based accounting? *Current Opinion in Environmental Sustainability,* 2 (4): 225-230.

Ciais, P., P.P. Tans, M. Trolier, J.W.C. White, & R.J. Francey (1995), A large northern hemisphere terrestrial CO_2 sink indicated by the $^{13}C/^{12}C$ ratio of atmospheric CO_2. *Science*, 269(5227): 1098-1102.

Cook, R., D.N. Huntzinger, M. Post, A. Jacobson, Y. Wei, and NACP Interim Synthesis Participants (2009), Preliminary results of the NACP Regional Interim Synthesis, presentation at AmeriFlux Annual Meeting, Washington, D.C., Sep. 22, 2009, available at http://nacp.ornl.gov/int_synth_contreg.shtml.

Corbin, K.D., A.S. Denning, E.Y. Lokupitiya, A.E. Schuh, N.L. Miles, K.J. Davis, S. Richardson, and I.T. Baker (2010), Assessing the impact of crops on regional CO_2 fluxes and atmospheric concentrations, *Tellus B*, 62, 5, Sp. Iss. SI, 521-532.

Cramer, W., A. Bondeau, F.I. Woodward, I.C. Prentice, R.A. Betts, V. Brovkin, P.M. Cox, V. Fisher, J.A. Foley, A.D. Friend, C. Kucharik, M.R. Lomas, N. Ramankutty, S. Sitch, B. Smith, A. White, and C. Young-Molling (2001), Global response of terrestrial ecosystem structure and function to CO_2 and climate change: results from six dynamic global vegetation models, *Global Change Biology*, 7, 4, 357-373.

Cramer, W., D. W. Kicklighter, A. Bondeau, B. Moore, C. Churkina, B. Nemry, A. Ruimy, and A. L. Schloss (1999), Comparing global models of terrestrial net primary productivity (NPP): overview and key results, *Global Change Biology*, *5*, 1-15.

Cressie, N. A. C. (1993), Statistics for Spatial Data, 900 pp., John Wiley, New York.

Crevoisier, C., C. Sweeney, M. Gloor, J. L. Sarmiento, and P. P. Tans (2010), Regional US carbon sinks from three-dimensional atmospheric CO_2 sampling, *Proceedings of the National Academy of Sciences of the United States of America*, 107(43), 18348–18353.

Dai, A. G., K. E. Trenberth, and T. T. Qian (2004), A global dataset of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming, *J Hydrometeorol, 5*(6), 1117-1130.

Dargaville, R., A.D. McGuire, and P. Rayner (2002), Estimates of large-scale fluxes in high latitudes from terrestrial biosphere models and an inversion of atmospheri CO_2 measurements, *Climatic Change*, 55, 1-2, 273-285.

Deng, F., J.M. Chen, M. Ishizawa, C.W. Yuen, G. Mo, K. Higuchi, D. Chan, and S. Maksyutov (2007), Global monthly CO_2 flux inversion with a focus over North America, *Tellus B*, 59, 2, 179-190.

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. Zhang, D.A. Randall, J.A. Berry, P.J. Sellers, G.D. Colello, & D.A. Dazlich (1996), Simulations of terrestrial carbon metabolism and atmospheric CO_2 in a general circulation model: 1. Surface carbon fluxes, *Tellus*, 48B, 521-542.

Denning A.S., N. Zhang, CX Yi, M. Branson, K. Davis, J. Kleist, and P. Bakwin (2008), Evaluation of modeled atmospheric boundary layer depth at the WLEF tower, *Agr. For. Met.*, 148, 2, 206-215.

Desai, A. R., B. R. Helliker, P. R. Moorcroft, A. E. Andrews, and J. A. Berry (2010), Climatic controls of interannual variability in regional carbon fluxes from top-down and bottom-up perspectives, *Journal of Geophysical Research-Atmospheres,* 115, G02011, doi:10.1029/2009JG001122.

Djuricin, S., D.E. Pataki, and X.M. Xu (2010), A comparison of tracer methods for quantifying CO_2 sources in an urban region, *J. Geophys. Res.-Atm.*, 115, D11303.

Ek, M.B., K.E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J.D. Tarpley (2003), Implementation of Noah land surface model advances in the National Centers for Enviromental Prediction operational mesoscale Eta model, *J. Geophys. Res.-Atm.*, 108, D22, 8851.

Energy Information Administration, U. S. Department of Energy (2002), Table H.1: World carbon dioxide emissions from the consumption and flaring of fossil fuels, 1980-- present, *Int. Energy Annu.*, 2002. (Available at http://www.eia.doe.gov/emeu/iea/tableh1.html).

Engelen, R. J., A. S. Denning, and K. R. Gurney (2002), On error estimation in atmospheric CO_2 inversions*, J. Geophys. Res.,* 107(D22), 4635, doi:10.1029/2002JD002195.

Enting, I. G. (2002), Inverse Problems in Atmospheric Constituent Transport, Cambridge Univ. Press, Cambridge, U. K. Enting, I. G., and G. N. Newsam (1990), Atmospheric constituent inversion problems—Implications for base-line monitoring, J. Atmos. Chem., 11(1–2), 69– 87.

Fan, S., M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, & P. Tans (1998), A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science*, 282: 442-446.

Fan, S., T. L. Blaine, and J. L. Sarmiento (1999), Terrestrial carbon sink in the northern hemisphere estimated from atmospheric CO_2 difference between Mauna Loa and South Pole since 1959, *Tellus,* 51, 863–870.

Farquhar, G.D. & M.L. Roderick (2003), Atmospheric science: Pinatubo, diffuse light, and the carbon cycle, *Science*, 299, 5615, 1997-1998.

Finzi, A.C., D.J.P. Moore, E.H. DeLucia, J. Lichter, K.S. Hofmockel, R.B. Jackson, H.S. Kim, R. Matamala, H.R. McCarthy, R. Oren, J.S. Pippen, & W.H. Schlesinger (2006), Progressive nitrogen limitation of ecosystem processes under elevated CO_2 in a warm-temperate forest, *Ecology* 87(1): 15-25.

Forster, P., V. Ramaswamy, P. Artaxo, T. Berntsen, R. Betts, D.W. Fahey, J. Haywood, J. Lean, D.C. Lowe, G. Myhre, J. Nganga, R. Prinn, G. Raga, M. Schulz and R. Van Dorland (2007), Changes in Atmospheric Constituents and in Radiative Forcing. *In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Francey, R.J., C.M. Trudinger, M. van der Schoot, P.B. Krummel, L.P. Steele, and R.L. Langenfelds (2010), Differences between trends in atmospheric CO_2 and the reported trends in anthropogenic CO_2 emissions, *Tellus B*, 62, 5, Sp. Iss. SI, 316-328.

Francey, R.J. & L.P. Steele (2003), Measuring atmospheric carbon dioxide – the calibration challenge, *Accreditation and Quality Assurance*, 8, 5, 200-204.

Friedlingstein, P. & I.C. Prentice (2010), Carbon-climate feedbacks: a review of model and observation based estimates, *Current Opinion in Environmental Sustainability*, 2, 4, 251-257.

Friedlingstein, P., P. Cox, R. Betts, L. Bopp, W. Von Bloh, V. Brovkin, P. Cadule, S. Doney, M. Eby, I. Fung, G. Bala, J. John, C. Jones, F. Joos, T. Kato, M. Kawamiya, W. Knorr, K. Lindsay, H.D. Matthews, T. Raddatz, P. Rayner, C. Reick, E. Roeckner, K.G. Schnitzler, R. Schnur, K. Strassmann, A.J. Weaver, C. Yoshikawa, & N. Zeng (2006), Climate-carbon cycle feedback analysis: Results from the (CMIP)-M-4 model intercomparison. *Journal of Climate*, 19(14): 3337-3353.

Friend, A.D., A. Arneth, N.Y. Kiang, M. Lomas, J. Ogee, C. Rodenbeck, S.W. Running, J. Santaren, S. Sitch, N. Viovy, F.I Woodward, & S. Zaehle (2007), FLUXNET and modeling the global carbon cycle. *Global Change Biology*, 13, 610-633, doi:10.1111/ /j.1365-2486.2006.01223.x.

Geels, C., S.C. Doney, R. Dargaville, J. Brandt, and J.H. Christensen (2004), Investigating the sources of synoptic variability in atmospheric CO_2 measurements over the Northern Hemisphere continents: a regional model study, Tellus, 56B, 35-50.

Geels, C., M. Gloor, P. Ciais, P. Bousquet, P. Peylin, A.T. Vermeulen, R. Dargaville, T. Aalto, J. Brandt, J.H. Christensen, L.M. Frohn, L. Haszpra, U. Karstens, C. Rödenbeck, M. Ramonet, G. Carboni, and R. Santaguida (2007), Comparing atmospheric transport models for future regional inversions over Europe - Part 1: mapping the atmospheric CO_2 signals, Atm. Chem. Phys., 7, 3461-3479.

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 CO_2 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 CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework, *J. Geophys. Res.*, 108 (D24), 4757, doi:10.1029/2003JD003770.

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 fluxes?, *Atm. Chem. Phys.*, 6, 539-554.

Gerbig, C., S. Körner, & J.C. Lin (2008), Vertical mixing in atmospheric tracer transport models: error characterization and propagation. *Atm. Chem. Phys.*, 8, 591-602.

Gerbig, C., A.J. Dolman, and M. Heimann (2009), On observational and modeling strategies targeted at regional carbon exchange over continents. Biogeosciences, 6, 1949-1959.

GLOBALVIEW-CO_2 (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/CO_2/GLOBALVIEW).

Gloor, M. S-M. Fan, S. Pacala & J. Sarmiento (2000), Optimal sampling of the atmosphere for purpose of inverse modeling: A model study. *Global Biogeochemical Cycles*, 14, 1, 407-428.

Gloor, M., P. Bakwin, D. Hurst, L. Lock, R. Draxler, & P. Tans (2001), What is the concentration footprint of a tall tower? *J. Geophys. Res.*, 106, D16, 17831-17840.

Göckede, M., A. M. Michalak, D. Vickers, D. P. Turner, and B. E. Law (2010a), Atmospheric inverse modeling to constrain regional-scale CO_2 budgets at high spatial and temporal resolution, *J. Geophys. Res.*, 115, D15113, doi:10.1029/2009JD012257.

Göckede, M., D.P. Turner, A.M. Michalak, D. Vickers, and B.E. Law (2010b), Sensitivity of a sub-regional scale atmospheric inverse CO_2 modeling framework to boundary conditions, *J. Geophys. Res.*, 115, D24112, doi:10.1029/2010JD014443.

Gourdji, S.M., K.L. Mueller, K. Schaefer, & A.M. Michalak (2008), Global monthly-averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 2. Results including auxiliary environmental data, *J. Geophys. Res.*, 113(D21115), doi:10.1029/2007JD009733.

Gourdji, S.M., A.I. Hirsch, K.L. Mueller, V. Yadav, A.E. Andrews, & A.M. Michalak (2010), Regional-scale geostatistical inverse modeling of North American CO_2 fluxes: a synthetic data study, *Atm. Chem. Phys.*, 10, 6151-6167.

Grell, G.A. & D. Devenyi (2002), A generalized approach to parameterizing convection combining ensemble and data assimilation techniques, *Geophys. Res. Lett.*, 29, 14, 1693.

Gurney, K., R.M. Law, Denning A.S., Rayner P.J., Baker D., Bousquet P., Bruhwiler L., Chen Y.H., Ciais P., Fan S., Fung I.Y., Gloor M., Heimann M., Higuchi K., John J., Maki T., Maksyutov S., Masarie K., Peylin P., Prather M., Pak B.C., Randerson J., Sarmiento J., Taguchi S., Takahashi T. & Yuen C.W. (2002), Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature, 415(6872): 626-630.

Gurney, K. R., R. M. Law, A. S. Denning, P. J. Rayner, D. Baker, P. Bousquet, L. Bruhwiler, Y. H. Chen, P. Ciais, S. M. Fan, I. Y. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, E. Kowalczyk, T. Maki, S. Maksyutov, P. Peylin, M. Prather, B. C. Pak, J. Sarmiento, S. Taguchi, T. Takahashi, and C. W. Yuen (2003), TransCom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus B, 55*(2), 555-579.

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

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

Gurney, K.R., D. Baker, P. Rayner, and A.S. Denning (2008), Interannual variations in continental-scale net carbon exchange and sensitivity to observing networks estimated from atmospheric CO_2 inversions for the period 1980 to 2005, *Glob. Biogeoch. Cyc.*, 22, 3, GB3025.

Gurney, K. R., Mendoza, D. L., Zhou, Y., Fischer, M. L., Miller (2009), High resolution fossil fuel combustion CO_2 emission fluxes for the United States. *Environ. Sci. Technol.* 43, doi:10.1021/es900806c.

Harris, D.C. (2010), Charles David Keeling and the story of atmospheric CO_2 measurements, *Analytical Chemistry*, 82, 19, 7865-7870.

Haszpra, L. (1999), On the representativeness of carbon dioxide measurements, J. Geophys. Res., 104 (D21), 26,953-26,960.

Heimann, M., and S. Körner (2003), The Global Atmospheric Tracer Model TM3: Model Description and User's Manual, Release 3.8a, Jena.

Hoeksema, R.J. & P.K. Kitanidis (1984), An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resources Research*, 20, 7, 1003-1020.

Huntzinger, D.N., S.M. Gourdji, K.L. Mueller & A.M. Michalak (2011), "The utility of continuous atmospheric measurements for identifying biospheric CO_2 flux variability", *J.Geophys.Res.-Atm.,* doi:10.1029/2010JD015048, in press.

Huntzinger, D.N., S.M. Gourdji, K.L. Mueller & A.M. Michalak (2010), A quantitative approach for comparing modeled biospheric carbon flux estimates across regional scales, *Biogeosciences Discuss.*, 7, 7903-7943, doi:10.5194/bgd-7-7903-2010.

Huntzinger, D.N., W. Post, A.M. Michalak, Y. Wei, A. Jacobson, R. Cook and co-authors (in prep), North American Carbon Program (NACP) Regional Interim Synthesis: Terrestrial biospheric model inter-comparison.

Ise, T. and P.R. Moorcroft (2006), The global-scale temperature and moisture dependencies of soil organic carbon decomposition: an analysis using a mechanistic decomposition model, *Biogeochemistry*, 80, 3, 217-231.

Jain, A. K., and X. J. Yang (2005), Modeling the effects of two different land cover change data sets on the carbon stocks of plants and soils in concert with CO_2 and climate change, *Global Biogeochemical Cycles*, 19, 2, GB2015.

Jones, C.D. & P.J. Cox (2005), On the significance of atmospheric CO_2 growth rate anomalies in 2002-2003, *Geophys. Res. Lett.*, 32, 14, L14816.

Judd, C. M., and G. H. McClelland (1989), *Data Analysis: A Model-Comparison Approach*, xviii + 528 + 107 pp., Harcourt Brace Jovanovich, San Diego.

Jung, M., M. Reichstein, and A. Bondeau (2009), Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, 6, 10, 2001-2013.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, B. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society, 77*, 437-471.

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, *J Geophys Res-Atmos, 104*(D15), 18555-18581.

Kaminski, T., P.J. Rayner, M. Heimann, and I.G. Enting (2001), On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106, 4703-4715.

Keeling, C.D., T.P. Whorf, and J. Vanderplicht (1995), Interannual extremes in the rate of rise of atmospheric carbon-dioxide since 1980, *Nature*, 375,6533, 666-670.

Kelley, J.J., D.F. Weaver, and B.P. Smith (1968), Variation of carbon dioxide under snow in Arctic, *Ecology*, 49, 2, 358-361.

Kicklighter, D. W., A. Bondeau, A. L. Schloss, J. Kaduk, A. D. McGuire, and N. P. P. M. I. Participants Potsdam (1999), Comparing global models of terrestrial net primary productivity (NPP): global pattern and differentiation by major biomes, *Global Change Biology*, *5*, 16-24.

Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, *Water Resour Res, 31*(10), 2411-2419.

Kitanidis, P. K. (1997), A variance-ratio test for supporting a variable mean in kriging, *Mathematical Geology, 29*(3), 335-349.

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, GB1015.

Land, A.H., and A.G. Doig (1960), An automatic method of solving discrete programming problems, *Econometrica*, 28, 3, 497-520.

Lauvaux, T., M. Uliasz, C. Sarrat, F. Chevallier, P. Bousquet, C. Lac, K.J. Davis, P. Ciais, A.S. Denning, and P.J. Rayner (2008), Mesoscale inversion: first results from the CERES campaign with synthetic data, Atm. Chem. & Phys., 8(13), 3459-3471.

Law, R.M., P.J. Rayner, L.P. Steele, and I.G. Enting (2002), Using high temporal frequency data for CO_2 inversions, Global Biogeochem Cy, 4, 1053, doi:10.1029/2001GB001593.

Law, R. M., Y. H. Chen, and K. R. Gurney (2003a), TransCom 3 CO_2 inversion intercomparison: 2. Sensitivity of annual mean results to data choices, *Tellus*, Ser. B, 55, 580– 595.

Law, R.M., P.J. Rayner, L.P. Steele, and I.G. Enting (2003b), Data and modelling requirements for CO_2 inversions using high frequency data, Tellus, 55B, 512-521, doi: 10.1034/j.1600-0560.2003.0029.x.

Law, R.M., P.J. Rayner, and Y.P. Wang (2004), Inversion of diurnally-varying synthetic CO_2: network optimisation for an Australian test case, Global Biogeochem Cy, 18, GB1044, doi: 10.1029/2003GB002136.

Levin, I. & U. Karstens (2007), Inferring high-resolution fossil fuel CO_2 records at continental sites from combined $^{14}CO_2$ and CO observations. *Tellus*, 59B, 245-250.

Li, Y. F. (1996), Global Population Distribution Database, A Report to the United Nations Environment Programme, under UNEP Sub-Project FP/1205-95-12.

Lin, J.C. & C. Gerbig (2005), Accounting for the effect of transport errors on tracer inversions. *Geophysical Research Letters*, 32, L01802, doi:10.1029/2004GL021227.

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, J. Geophys. Res., 108 (D16), 4493, doi:10.1029/2002JD003161.

Lin, J.C., C. Gerbig, S.C. Wofsy, A.E. Andrews, B.C. Daube, C.A. Grainger, B.B. Stephens, P.S. Bakwin, and D.Y. Hollinger (2004), Measuring fluxes of trace gases at regional scales by Lagrangian observations: application to the CO_2 Budget and Rectification Airborne (COBRA) study, J. Geophys. Res., 109 D15304, doi:10.1029/2004JD004754).

Lloyd, J., and J. A. Taylor (1994), On the temperature-dependence of soil respiration, *Functional Ecology*, *8*(3), 315-323.

Lokupitiya, E., A.S. Denning, K. Paustian, I. Baker, K. Schaefer, S. Verma, T. Meyers, C.J. Bernacchi, A. Suyker, and M. Fischer (2009), Incorporation of crop phenology in Simple Biosphere Model (SiBCrop) to improve land-atmosphere carbon exchanges from croplands, *Biogeosciences*, 6, 6, 969-986.

Los, S. O., G. J. Collatz, L. Bounoua, P. J. Sellers, and C. J. Tucker (2001), Global interannual variations in sea surface temperature and land surface vegetation, air temperature, and precipitation, *Journal Of Climate, 14*(7), 1535-1549.

Loreau, M., N. Mouquet, and R.D. Holt (2003), Meta-ecosystems: a theoretical framework for a spatial ecosystem ecology, *Ecology Letters*, 6, 8, 673-679.

Loveland, T. R., B. C. Reed, J. F. Brown, D. O. Ohlen, J. Zhu, L. Yang, and J. W. Merchant (2001), Development of a Global Land Cover Characteristics Database and IGBP DISCover from 1-km AVHRR Data, *International Journal of Remote Sensing, 21*(6/7), 1303-1330.

Mahadevan, P., S. C. Wofsy, D. M. Matross, X. M. Xiao, A. L. Dunn, J. C. Lin, C. Gerbig, J. W. Munger, V. Y. Chow, and E. W. Gottlieb (2008), A satellite-based biosphere parameterization for net ecosystem CO_2 exchange: Vegetation Photosynthesis and Respiration Model (VPRM), *Global Biogeochem. Cycles*, 22, GB2005, doi:10.1029/2006GB002735.

Maki, T., M. Ikegami, T. Fujita, T. Hirahara, K. Yamada, K. Mori, A. Takeuchi, Y. Tsutsumi, K. Suda, and T.J. Conway (2010) New technique to analyse global distributions of CO_2 concentrations and fluxes from non-processed observational data, *Tellus*, *62B, 797–80,* DOI: 10.1111/j.1600-0889.2010.00488.x.

Marland, G., T. A. Boden, and R. J. Andres (2006), Global, Regional, and National Fossil Fuel CO_2 Emissions. In *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Marland G., K. Hamal, and M. Jonas (2009), How uncertain are estimates of CO_2 emissions?, *J. Ind. Ecol.,* 13, 1, 4-7.

Masarie, K.A. and P.P. Tans (1995), Extension and integration of atmospheric carbon-dioxide data into a globally consistent measurement record, J*. Geophys. Res.-Atm.*, 100, D6, 11593-11610.

Masera, O.R., M.J. Ordonez, & R. Dirzo (1997), Carbon emissions from Mexican forests: current situation and long-term scenarios. *Climatic Change*, 35(3), doi:10.1023/A:1005309908420.

MathWorks (2010), Trust-region reflective fsolve algorithm. Retrieved Jan. 23, 2011. http://www.mathworks.com/access/helpdesk/help/toolbox/optim/ug/brnoyhf.html

Matross, D.M., A. Andrews, M. Pathmathevan, C. Gerbig, J.C. Lin, S.C. Wofsy, B.C. Daube, E.W. Gottlieb, V.Y. Chow, J.T. Lee, C. Zhao, P.S. Bakwin, J.W. Munger, and D.Y. Hollinger (2006), Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data. *Tellus* B, 58: 344–358. doi: 10.1111/j.1600-0889.2006.00206.x

McGuire, A.D., D.J. Hayes, D. W. Kicklighter, M. Manizza, Q. Zhuang, M. Chen, M.J. Follows, K.R. Gurney, J.W. McLelland, J.M. Melillo, B.J. Peterson, and R.G. Prinn (2010), An analysis of the carbon balance of the Arctic Basin from 1997 to 2006, *Tellus B*, 62, 5, Sp. Iss. SI, 455-474.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, D. Jovic, J. Woolen, E. Rogers, E.H. Berbery, M.B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi (2006), North American Regional Reanalysis, *Bull. Am. Meteo. Soc.*, 87, 3, 343-360.

Michalak, A. M., L. Bruhwiler, and P. P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, *J Geophys Res-Atmos, 109*(D14109), 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, D24107, doi:10.1029/2005JD005970.

Mueller, K. L., V. Yadav, P. S. Curtis, C. Vogel, and A. M. Michalak (2010), Attributing the variability of eddy-covariance CO_2 flux measurements across temporal scales using geostatistical regression for a mixed northern hardwood forest, *Global Biogeochem. Cycles*, 24, GB3023, doi:10.1029/2009GB003642.

Mueller, K.L., S.M. Gourdji, & A.M. Michalak (2008), Global monthly-averaged CO_2 fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements. *J. Geophys. Res.*, 113(D21114), doi:10.1029/2007JD009734.

Mueller, K.L., S.M. Gourdji, V. Yadav, M. Trudeau, A. Chatterjee, D. N. Huntzinger, A.E. Andrews, A. Schuh, Y. Shiga, K. J. Davis, B.B. Stephens, B.E. Law, C. Sweeney, M. Fischer, D. Dragoni, D. Worthy, M. Parker, A. M. Michalak (in prep.), Impact of the expanding measurement network on top-down budgeting of CO_2 surface fluxes in North America.

National Research Council (2010), *Verifying Greenhouse Gas Emissions: Methods to Support International Climate Agreements*, Committee on Methods for Estimating Greenhouse Gas Emissions, ISBN: 0-309-15212-7, 124 pp.

Nehrkorn, T., J. Eluszkiewicz, S.C. Wofsy, J.C. Lin, C. Gerbig, M. Longo, S. Freitas (2010), Coupled weather research and forecasting–stochastic time-inverted lagrangian transport (WRF–STILT) model, *Meteorol Atmos Phys*, 107:51–64.doi:10.1007/s00703-010-0068-x.

Nemani, R., M. White, P. Thornton, K. Nishida, S. Reddy, J. Jenkins & S. Running (2002), Recent trends in hydrologic balance have enhanced the terrestrial carbon sink in the United States. *Geophysical Research Letters*, 29(10), 1468, 10.1029/2002GL014867.

Norby, R.J., J.M. Warren, C.M. Iversen, B.E. Medlyn, and R.E. McMurtrie (2010), CO_2 enhancement of forest productivity constrained by limited nitrogen availability, *Proc. Natl. Acad. Sci.*, 107, 45, 19368-19373.

Oda, T. & S. Maksyutov (2010), A very high-resolution global fossil fuel CO_2 emission inventory derived using a point source database and satellite observations of night-time lights, 1980-2007, *Atm. Chem. Phys. Disc.*, 10, 16307-16344.

Olsen, J.C. and J.T. Randerson (2004), Differences between surface and column atmospheric CO_2 and implications for carbon cycle research, J. Geophys. Res.-Atm., 109, D02301, doi:10.1029/2003JD003968.

Olson, D.M., E. Dinerstein, E.D. Wikramanayake, N.D. Burgess, G.V.N. Powell, E.C. Underwood, J.A. D'Amico, I. Itoua, H.E. Strand, J.C. Morrison, C.J. Loucks, T.F. Allnutt, T.H. Ricketts, Y.Kura, J.F. Lamoreux, W.W. Wettengel, P. Hedao, and K.R. Kassem (2001), Terrestrial ecoregions of the world: a new map of life on earth, Bioscience, 51(11), 933-938.

Pacala, S. & R. Socolow (2004), Stabilization wedges: solving the climate problem for the next 50 years with current technologies, *Science*, 305, 5686, 968-972.

Pacala, S. W., et al. (2001), Consistent land- and atmosphere-based US carbon sink estimates, *Science*, *292*(5525), 2316-2320.

Palmer, W. C. (1965), Meteorological drought, Research Paper 45, U. S. Department of Commerce, pp. 58.

Pataki, D.E., D.R. Bowling, J.R. Ehleringer, & J.M. Zobitz (2006a), High resolution atmospheric monitoring of urban carbon dioxide sources. *Geophysical Research Letters*, 33, L03813, doi:10.1029/2005GL024822.

Patra, P.K., R.M. Law, W. Peters, C. Rödenbeck, M. Takigawa, C. Aulagnier, I. Baker, D.J. Bergmann, P. Bousquet, J. Brandt, L. Bruhwiler, P.J. Cameron-Smith, J.H. Christensen, F. Delage, A.S. Denning, S. Fan, C. Geels, S. Houweling, R. Imasu, U. Karstens, S.R. Kawa, J. Kleist, M.C. Krol, S.-J. Lin, R.

Lokupitiya, T. Maki, S. Maksyutov, Y. Niwa, R. Onishi, N. Parazoo, G. Pieterse, L. Rivier, M. Satoh, S. Serrar, S. Taguchi, R. Vautard, A.T. Vermeulen, and Z. Zhu (2008), TransCom model simulationsof hourly atmospheric CO_2: Analysis of synoptic-scale variations for the period 2002-2003, Global Biogeochem Cy, 22, GB4013, doi:10.1029/2007GB003081.

Patra, P., K.R. Gurney, A.S. Denning, S. Maksyutov, T. Nakazawa, D. Baker, P. Bousquet, L. Bruhwiler, Y.H. Chen, P. Ciais, S.M. Fan, I. Fung, M. Gloor, M. Heimann, K. Higuchi, J. John, R.M. Law, T. Maki, B.C. Pak, P. Peylin, M. Prather, P.J. Rayner, J. Sarmiento, S. Taguchi, T. Takahashi, and C.W. Yuen (2006), Sensitivity of inverse estimation of annual mean CO_2 sources and sinks to ocean-only sites vs. all-sites observational networks, *Geophysical Res. Lett.*, 33, 5, L05814.

Peters, W., J.B. Miller, K.M. Schaefer, I. van der Velde, G. van der Werf, A.J. Dolman, N. Carvalhais, and P.P. Tans (2010), Forest carbon imbalance information improves atmosphere based carbon data assimilation systems, abstract B31F-0378 presented at 2010 Fall Meeting, AGU, San Francisco, CA, 13-17 Dec.

Peters, W., A.R. Jacobson, C. Sweeney, A.E. Andrews, T.J. Conway, K. Masarie, J.B. Miller, L.M.P. Bruhwiler, G. Petron, A.I. Hirsch, D.E.J. Worthy, G.R. van der Werf, J.T. Randerson, P.O. Wennberg, M.C. Krol & P.P. Tans (2007), An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, PNAS, 104(48).

Peylin, P., P.J. Rayner, P. Bousquet, C. Carouge, F. Hourdin, P. Heinrich, P. Ciais, and AEROCARB contributors (2005), Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atm. Chem. Phys., 5, 3173-3186.

Peylin P., D. Baker, J. Sarmiento, P. Ciais, and P. Bousquet (2002), Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO_2 data, *J. Geophys. Res.-Atm.*, 107, D19, 4385.

Pisso, I.J., P.K. Patra, T. Nakazawa, Y. Sawa, T. Machida, and H. Matsueda (2010), Constraints on CO_2 flux emissions: reconstructions of in-situ measurements from Lagrangian stochastic inversion, *Eos Trans. AGU*, 91(26), Meet. Am. Suppl., Abstract B31B-06.

Potter, C., P. Gross, S. Klooster, M. Fladeland, and V. Genovese (2008), Storage of carbon in US forests predicted from satellite data, ecosystem modeling, and inventory summaries, *Climatic Change*, 90, 3, 269-282.

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), 811-841.

Prather, M.J., X. Zhua, S.E. Strahan, S.D. Steenrod, and J.M. Rodriguez (2008), Quantifying errors in trace species transport modeling, *Proc. Natl. Acad. Sci.*, 105, 50, 19617-19621.

Pregitzer, K.S., A.J. Burton, D.R. Zak, & A.F. Talhelm (2008), Simulated chronic nitrogen deposition increases carbon storage in Northern Temperate forests, *Global Change Biology*, 14(1): 142-153.

Ramonet & Monfray (1996), CO_2 baseline concept in 3-D atmospheric transport models, *Tellus B*, 48, 4, 502-520.

Randerson, J.T., F.M. Hoffman, P.E. Thornton, N.M. Mahowald, K. Lindsay, Y.H. Lee, C.D. Nevison, S.C. Doney, G. Bonan, R. Stockli, C. Covey, S.W. Running, and I.Y. Fung (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biology*, 15, 10, 2462-2484.

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.

Rayner, P. J., I. G. Enting, R. J. Francey, and R. Langenfelds (1999), Reconstructing the recent carbon cycle from atmospheric CO_2, delta C-13 and O-2/N-2 observations, *Tellus*, Ser. B, 51, 213–232.

Rayner, P.J., M. Scholze, T. Kaminski, R. Giering, and H. Widmann (2005), Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), *Glob. Biogeo. Cy.*, 19, 2, GB2026.

Rayner, P.J., R.M. Law, C.E. Allison, R.J. Francey, C.M. Trudinger, & 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, GB3008, doi:10.1029/2007GB003068.

Reichstein, M., A. Rey, A. Freibauer, J. Tenhunen, R. Valentini, J. Banza, P. Casals, Y. F. Cheng, J. M. Grunzweig, J. Irvine, R. Joffre, B. E. Law, D. Loustau, F. Miglietta, W. Oechel, J. M. Ourcival, J. S. Pereira, A. Peressotti, F. Ponti, Y. Qi, S. Rambal, M. Rayment, J. Romanya, F. Rossi, V. Tedeschi, G. Tirone, M. Xu, and D. Yakir (2003), Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices, *Global Biogeochemical Cycles, 17*(4): Art. No. 1104, doi: 10.1029/2003GB002035.

Rodell, M., P.R. Houser, U. Jambor, J. Gottschalck, K. Mitchell, C.J. Meng, K. Arsenault, B. Cosgrove, J. Radakovich, M. Bosilovich, J.K. Entin, J.P. Walker, D. Lohmann, and D. Toll (2004), The global land data assimilation system, *Bulletin of the American Meteorological Society*, 85(3): 381-394.

Rödenbeck, C.., C. Gerbig, K. Trusilova, and M. Heimann (2009), A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, *Atm. Chem. Phys.*, 9, 14, 5331-5342.

Rödenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO_2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos Chem Phys*, *3*, 1919-1964.

Sarmiento J.L., and N. Gruber (2006), *Ocean Biogeochemical Dynamics*, Princeton University Press, Princeton, N.J.

Sarrat, C., J. Noilhan, A.J. Dolman, C. Gerbig, R. Ahmadov, L.F. Tolk, A.G.C.A. Meesters, R.W.A. Hutjes, H.W. Ter Maat, G. Perez-Landa, and S. Donier (2007), Atmospheric CO_2 modeling at the regional scale: an intercomparison of 5 meso-scale atmospheric models, *Biogeosciences*, 4, 6, 1115-1126.

Schaefer, K., A. S. Denning, N. Suits, J. Kaduk, I. Baker, S. Los, and L. Prihodko (2002), Effect of climate on interannual variability of terrestrial CO_2 fluxes, *Global Biogeochemical Cycles, 16*(4): 1102, doi: 10.1029/2002GB001928.

Schaefer, K., A. S. Denning, and O. Leonard (2005), The winter Arctic Oscillation, the timing of spring, and carbon fluxes in the Northern Hemisphere, *Global Biogeochemical Cycles, 19*(GB3017), doi: 10.1029/2004GB002336.

Scholze, M., T. Kaminski, P. Rayner, W. Knorr, and R. Giering (2007), Propagating uncertainty through prognostic carbon cycle data assimilation system simulations, *J. Geophys. Res. – Atm.*, 112, D17, D17305.

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

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), 1625-1644, doi: 10.5194/bg-7-1625-2010.

Schwalm, C.R., C.A. Wiliams, K. Schaefer, R. Anderson, M.A. Arain, I. Baker, A. Barr, T.A. Black, G.S. Chen, J.M. Chen, P. Ciais, K.J. Davis, A. Desai, M. Dietze, D. Dragoni, M.L. Fischer, L.B. Flanagan, R. Grant, L.H. Gu, D. Hollinger, R.C. Izaurralde, C. Kucharik, P. Lafleur, B.E. Law, H. Li, Z.P. Li, S.G. Liu, E. Lokupitiya, Y.Q. Luo, S.Y. Ma, H. Margolis, R. Matamala, H. McCaughey, R.K. Monson, W.C. Oechel, C.H. Peng, B. Poulter, D.T. Price, D.M. Ricciuto, W. Riley, A.K. Sahoo, M. Sprintsin, J.F. Sun, H.Q. Tian, C. Tonitto, H. Verbeeck, and S.B. Verma (2010), A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis, *J. Geophys. Res.-Biogeo.*, 115, G00H05.

Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.,* 6(2),461–464, doi:10.1214/aos/1176344136.

Searchinger, T., R. Heimlich, R.A. Houghton, F. Dong, A. Elobeid, J. Fabiosa, S. Tokgoz, D. Hayes, & T.-H. Yu (2008), Use of U.S. Croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, 319, 1238-1240.

Sellers, P.J., D.A. Randall, G.J. Collatz et al. (1996a), A revised land-surface parameterization (SiB2) for atmospheric GCM's. Part 1: Model formulation. *Journal of Climate*, 1, 676-705.

Sellers, P. J., S. O. Los, C. J. Tucker, C. O. Justice, D. A. Dazlich, G. J. Collatz, and D. A. Randall (1996b), A revised land surface parameterization (SiB2) for atmospheric GCMs .2. The generation of global fields of terrestrial biophysical parameters from satellite data, *Journal Of Climate, 9*(4), 706-737.

Siegenthaler, U., T.F. Stocker, E. Monnin, D. Luthi, J. Schwander, B. Stauffer, D. Raynaud, J. Barnola, H. Fischer, V. Masson-Delmotte & J. Jouzel (2005), Stable carbon cycle-climate relationship during the late Pleistocene. *Science*, 310, 1313-1317.

Sitch, S., C. Huntingford, N. Gedney, P.E. Levy, M. Lomas, S.L. Piao, R. Betts, P. Ciais, P. Cox, P. Friedlingstein, C.D. Jones, I.C. Prentice, and F.I. Woodward (2008), Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVM's), *Global Change Biology*, 14, 9, 2015-2039.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, & J. G. Powers (2005), A description of the advanced research WRF version 2. Technical Note 468+STR, MMM Division, NCAR, Boulder, CO, 88 pp. Available from http://www.mmm.ucar.edu/wrf/users/docs/arw.pdf.

Stephens, B.B., K.R. Gurney, P.P. Tans, C. Sweeney, W. Peters, L. Bruhwiler, P. Ciais, M. Ramonet, P. Bousquet, T. Nakazawa, S. Aoki, T. Machida, G. Inoue, N. Vinnichenko, J. Lloyd, A. Jordan, M. Heimann, O. Shibistova, R.L. Langenfelds, L.P. Steele, R.J. Francey, and A.S. Denning (2007), Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, *Science*, 316, 5832, 1732-1735.

Stoy, P. C., et al. (2009), Biosphere-atmosphere exchange of CO_2 in relation to climate: A cross-biome analysis across multiple time scales, *Biogeosciences*, 6(10), 2297–2312, doi:10.5194/bg-6-2297-2009.

Takahashi, T., S. C. Sutherland, C. Sweeney, A. Poisson, N. Metzl, B. Tilbrook, N. Bates, R. Wanninkhof, R. A. Feely, C. Sabine, J. Olafsson, and Y. Nojiri (2002), Global sea-air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects, *Deep-Sea Res Pt II, 49*(9-10), 1601-1622.

Takahashi, T., S.C. Sutherland, R. Wanninkhof, C. Sweeney, R.A. Feely, D.W. Chipman, B. Hales, G. Friederich, F. Chavez, C. Sabine, A. Watson, D.C.E. Bakker, U. Schuster, N. Metzl, H. Yoshikawa-Inoue, M. Ishii, T. Midorikawa, Y. Nojiri, A. Kortzinger, T. Steinhoff, M. Hoppema, J. Olafsson, T.S. Arnarson, B. Tilbrook, T. Johannessen, A. Olsen, R. Bellerby, C.S. Wong, B. Delille, N.R. Bates, and H.J.W. de Baar (2009), Climatological mean and decadal change in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, *Deep-Sea Research II*, 56, 554-577.

Tans, P., and T. J. Conway (2005), Monthly Atmospheric CO_2 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.

Tans, P.P., I.Y. Fung, T. Takahashi (1990), Observational constraints on the global atmospheric CO_2 budget. *Science*, 247(4949):1431-1438.

Tolk, L.F., A.G.C.A. Meesters, A.J. Dolman, and W. Peters (2008), Modeling representation errors of atmospheric CO_2 mixing ratios at a regional scale, *Atm. Chem. Phys.*, 8, 22, 6587-6596.

Trusilova, K., C. Rodenbeck, C. Gerbig, and M. Heimann (2010), Technical note: a new coupled system for global-to-regional downscaling of CO_2 concentration estimation, *Atm. Chem. Phys.*, 10, 7, 3205-3213.

Tucker, C. J., and P. J. Sellers (1986), Satellite Remote-Sensing of Primary Production, *International Journal of Remote Sensing, 7*(11), 1395-1416.

Tucker, C. J., J. E. Pinzon, M. E. Brown, D. A. Slayback, E. W. Pak, R. Mahoney, E. F. Vermote, and N. El Saleous (2005), An Extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *International Journal of Remote Sensing, 26*(20), 4485-4498.

Uliasz, M. & R.A. Pielke (1991), Application of the receptor oriented approach in mesoscale dispersion modeling, *Air Pollution Modeling and its Applications VIII, NATO – Challenges of Modern Society*, 15, 399-407.

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, Atm. Chem. Phys., 6, 3423-3441.

Ward, E. J. (2008), A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools, *Ecol. Modell.,* 211(1–2), 1–10, doi:10.1016/j.ecolmodel.2007.10.030.

Welp, L.R., J.T. Randerson, & H.P. Liu (2007), The sensitivity of carbon fluxes to spring warming and summer drought depends on plant functional type in boreal forest ecosystems. *Agricultural and Forest Meteorology*, 147, 172-185.

West T.O., C.C. Brandt, L.M. Baskaran, C.M. Hellwinckel, R. Mueller, C.J. Bernacchi, V. Bandaru, B. Yang, B.S. Wilson, G. Marland, R.G. Nelson, D.G.D. Ugarte, and W.M. Post (2010), Cropland carbon fluxes in the United States: increasing geospatial resolution of inventory-based carbon accounting, *Ecological Applications*, 20, 4, 1074-1086.

West, G.B., B.J. Enquist, and J.H. Brown (2009), A general quantitative theory of forest structure and dynamics, *Proc. Natl. Acad. Sci.*, 106, 17, 7040-7045.

Xiao, J. F., et al. (2008), Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data, *Agricultural and Forest Meteorology*, *148*(11), 1827-1847.

Xiao, J. & A. Moody (2004), Photosynthetic activity of US biomes: responses to the spatial variability and seasonality of precipitation and temperature. *Global Change Biology*, 10, 437-451, doi: 10.1111/j.1529-8817.2003.00745.x.

Yadav, V., K.L. Mueller, & A.M. Michalak (in prep.), A heuristic and non-heuristic dual criteria discrete optimization algorithm for variable selection in spatial regression models.

Yadav, V., K. L. Mueller, D. Dragoni, A. M. Michalak (2010), A geostatistical synthesis study of factors affecting gross primary productivity in various ecosystems of North America, *Biogeosciences*, 7, 1445–1487, doi:10.5194/bgd-7-1445-2010.

Velazquez A., J.F. Mas, G. Bocco, and J.L. Palacio-Prieto (2010), Mapping land cover changes in Mexico, 1976-2000, and applications for guiding environmental management policy, *Singapore Journal of Tropical Geography*, 31, 2, 152-162.

Yang, W., N.V. Shabanov, D. Huang, W. Wang, R.E. Dickinson, R.R. Nemani, Y. Knyazikhin, and R.B. Myneni (2006), Analysis of leaf area index products from combination of MODIS Terra and Aqua data, *Remote Sensing of Environment*, 104, 3, 297-312.

Yetman, G., S. Gaffin, and D. Balk (2004), Global Grids of Gross Domestic Product (GDP) Density for 1990 [DVD/CD-ROM], ISLSCP Initiative II, NASA.

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

Zimmerman, D.A., G. de Marsily, C.A. Gotway, M.G. Marietta, C.L. Axness, R.L. Beauheim, R.L. Bras, J. Carrera, G. Dagan, P.B. Davies, D.P. Gallegos, A. Galli, J. Gomez-Hernandez, P. Grindrod, A.L. Gutjahr, P.K. Kitanidis, A.M. Lavenue, D. McLaughlin, S.P. Neuman, B.S. RamaRao, C. Ravenne, and Y. Rubin (1998), A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow, *Water Res. Res.*, 34, 6, 1373-1413.