Doctoral Thesis

Inverse modeling of the sources and sinks of atmospheric CO₂ joint constraints from the ocean and atmosphere

Author(s):
Steinkamp, Kay

Publication Date:
2011

Permanent Link:
https://doi.org/10.3929/ethz-a-007134329

Rights / License:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection. For more information please consult the Terms of use.
Abstract

This thesis addresses the quantification of carbon dioxide (CO$_2$) exchange between the three major compartments of the earth system: the atmosphere, the land and the ocean. Knowledge about the cycling of carbon among these compartments is crucial, since CO$_2$ is a greenhouse gas and its accumulation in the atmosphere in response to human activities is the primary driver of climate change. Inverse modeling techniques are applied to estimate the carbon exchange, in particular the so-called Bayesian synthesis inversion. This method uses observed concentrations of CO$_2$ and other carbon compounds in the atmosphere and the ocean in conjunction with models of atmospheric and oceanic tracer transport. The observed concentrations are propagated inversely (“back in time”) through the transport models to reconstruct the distribution of CO$_2$ sinks and sources at the atmosphere-surface interface. Sinks and sources (generic: fluxes) are resolved at continental scales and with monthly resolution averaged over time periods of 5 to 10 years between 1980 and 2008, allowing also for the estimation of decadal flux trends. Utilized observational data streams include atmospheric CO$_2$ concentrations from a network of surface stations, concentrations of dissolved inorganic carbon (DIC) in the ocean interior, and differences in the partial pressure of CO$_2$ (pCO$_2$) in the surface ocean and overlying atmospheric layer. Estimated CO$_2$ fluxes and trends are jointly constrained by, and simultaneously consistent with, all data streams. The joint approach allows to address persistent key questions in the field of CO$_2$ flux estimation, such as the strength of the carbon sink in Northern Hemisphere extra-tropical land regions in relation to the tropics: it is shown that the tropics are required to act as net source of carbon to the atmosphere in order to achieve consistency with the observations. A strong net CO$_2$ release from the Amazonian region plays a key role in this regard, especially as it has been persistent over the last three decades and appears robust across several sensitivity experiments. Most of the decadal increase in northern land CO$_2$ uptake is shown to be localized in boreal regions and can partially be attributed to enhanced growing season uptake.
Zusammenfassung

# Contents

1 Introduction 1

1.1 The global carbon cycle in the Anthropocene 1

  1.1.1 Atmospheric CO$_2$ and its role in climate change 2

  1.1.2 Global carbon budget 6

  1.1.3 Mechanisms behind the oceanic carbon sink 8

  1.1.4 Mechanisms behind the land carbon sink 10

1.2 Net carbon fluxes 13

  1.2.1 Sources of carbon 14

  1.2.2 Sinks for carbon 15

1.3 Profiling this Ph.D. project 21

2 Inverse method for carbon exchange across Earth’s surface 27

2.1 Data assimilation and Bayesian concept 27

2.2 Atmosphere-only CO$_2$ inversion 31

  2.2.1 Linear synthesis inversion 31

  2.2.2 Network choice and data handling 35

  2.2.3 Background flux fields 37

  2.2.4 Global CO$_2$ growth constraint 39

2.3 Ocean interior DIC inversion 40

  2.3.1 Preindustrial and anthropogenic carbon tracers 42

  2.3.2 Net air-sea flux 45

2.4 Surface ocean pCO$_2$-based flux information 46

2.5 Joint ocean-atmosphere inversion 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.1 Bayesian synthesis inversion</td>
<td>51</td>
</tr>
<tr>
<td>2.5.2 Inclusion of further constraints</td>
<td>55</td>
</tr>
<tr>
<td>2.6 Pseudo-inversion of CO$_2$ column data</td>
<td>57</td>
</tr>
<tr>
<td>2.6.1 Retrieval of CO$_2$ column data</td>
<td>57</td>
</tr>
<tr>
<td>2.6.2 Setup of pseudo-inversion</td>
<td>59</td>
</tr>
<tr>
<td>3 A multiple-constraint inversion for the estimation of seasonal</td>
<td>63</td>
</tr>
<tr>
<td>carbon sources and sinks</td>
<td></td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>63</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>64</td>
</tr>
<tr>
<td>3.3 Methods</td>
<td>66</td>
</tr>
<tr>
<td>3.3.1 Unregularized inversion of atmospheric CO$_2$</td>
<td>67</td>
</tr>
<tr>
<td>3.3.2 Ocean constraints</td>
<td>69</td>
</tr>
<tr>
<td>3.3.3 Land constraints</td>
<td>70</td>
</tr>
<tr>
<td>3.3.4 Selection of atmospheric transport models</td>
<td>71</td>
</tr>
<tr>
<td>3.3.5 Joint Inversion</td>
<td>71</td>
</tr>
<tr>
<td>3.4 Results and Discussion</td>
<td>73</td>
</tr>
<tr>
<td>3.4.1 Sea-to-air-fluxes</td>
<td>74</td>
</tr>
<tr>
<td>3.4.2 Northern extra-tropical land</td>
<td>76</td>
</tr>
<tr>
<td>3.4.3 Tropical land</td>
<td>79</td>
</tr>
<tr>
<td>3.4.4 Interrelation of Tropical source and Northern sink</td>
<td>83</td>
</tr>
<tr>
<td>3.4.5 Implications for the Amazonian carbon balance</td>
<td>84</td>
</tr>
<tr>
<td>3.4.6 Aggregated Tropical and Southern Land</td>
<td>86</td>
</tr>
<tr>
<td>3.4.7 Constraints from atmospheric CO$_2$ column pseudodata</td>
<td>89</td>
</tr>
<tr>
<td>3.5 Conclusions</td>
<td>91</td>
</tr>
<tr>
<td>4 Decadal trends of regional carbon fluxes from a joint ocean-</td>
<td>95</td>
</tr>
<tr>
<td>atmosphere inversion</td>
<td></td>
</tr>
<tr>
<td>4.1 Abstract</td>
<td>95</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>96</td>
</tr>
<tr>
<td>4.3 Methods</td>
<td>99</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 The global carbon cycle in the Anthropocene

Carbon dioxide (CO\textsubscript{2}) is among the best-studied trace gases in the atmosphere, due to its tight linkage to every biogeochemical process within the earth system as well as its importance for present-day and future climate change. The term "earth system" refers to the complex system comprising the atmosphere, hydrosphere, lithosphere, biosphere and the anthroposphere. The anthroposphere is the sphere of human activities interacting with the other spheres within the earth system. On geological time scales the anthroposphere is a very young player in the earth system, because humankind started to significantly influence the natural environment only a few centuries ago, coinciding with the onset of industrial revolution around 1750. The modifications imposed by humans on the earth system include changes to the global cycles of carbon, water and nutrients, together with climate, land cover and biodiversity (Crutzen, 2002; Steffen et al., 2004). Because the magnitude and the rate of change of the human perturbation are so great, the epoch since the onset of the industrial revolution is often called the "Anthropocene" to distinguish it from the Holocene that started about 12,000 years ago, despite its geological youth. Anthropocene trends are evident in the global carbon cycle, most directly through CO\textsubscript{2} emissions from fossil fuel combustion and other industrial processes, together with emissions from land-use change (LUC), such as deforestation. These emissions directly increase atmospheric CO\textsubscript{2} concentrations, which then propagate into the other compartments of the earth system (especially the oceans and terrestrial biosphere), where they perturb the local carbon cycles. Indirectly, the excess atmospheric CO\textsubscript{2} leads to current and future climate changes through its contribution to the atmospheric greenhouse effect.
1.1.1 Atmospheric CO\textsubscript{2} and its role in climate change

High-precision measurements of atmospheric CO\textsubscript{2} are essential to the understanding of the carbon cycle, because they tell us how much of the emitted CO\textsubscript{2} actually ends up in the atmosphere and, therefore, put constraints on the net carbon exchange between the earth system's compartments. In 1958, C.D. Keeling began the first systematic monitoring of CO\textsubscript{2} concentrations at Mauna Loa, Hawaii (approximately at 19\textdegree\textsc{N}), and at the South Pole (Keeling et al., 2001). He achieved high precision data with his implementation of a non-dispersive infrared gas analyzer. The precision was high enough to not only detect for the first time that CO\textsubscript{2} was increasing in the atmosphere, but also to resolve the seasonal cycle caused by seasonally varying photosynthesis and respiration in the terrestrial biosphere. In the inset of figure 1.1 part of his record of CO\textsubscript{2} concentration at Mauna Loa is shown.

Choosing remote sites such as Mauna Loa or the South Pole is based on the consideration that air sampled at such locations shows little short-term variation caused by local sources and sinks of CO\textsubscript{2}. Because CO\textsubscript{2} is a long-lived tracer in the atmosphere without internal atmospheric sources or sinks and because of the vigorous atmospheric transport and mixing
(time scales of a few months), \(\text{CO}_2\) is usually well-mixed in the atmosphere (IPCC, 2007a). This enables measurements made at such remote sites to provide an integrated picture of large parts of the surrounding surface including continents and city point sources. Measurements from only one or two of such sites are sufficient to constrain the global atmospheric \(\text{CO}_2\) evolution. In this thesis I make frequent use of this property (see chapters 3, 4 and figure 4.1). However, it was recognized in the 1980s and 1990s that if one aims to estimate \(\text{CO}_2\) sources and sinks over continental- or basin-scale areas, one would have to extend the \(\text{CO}_2\) measurement network. Yet, continuous \(\text{CO}_2\) analyzers are relatively expensive to maintain and require frequent calibration, putting limits to the extension of continuous observation networks. Instead, atmospheric \(\text{CO}_2\) measurements are now widely undertaken within air sample flask programs, where air is collected in glass or metal containers and subsequently sent to central well-calibrated laboratories for analysis. The most extensive international air sampling network is operated by the National Oceanic and Atmospheric Administration's Global Monitoring Division (NOAA/GMD) in the USA. It comprises measurements from continuous analyzer locations as well as flask air samples from a global network of roughly 100 sites, usually with weekly flask analyses.

This network can be further extended by the inclusion of non-surface \(\text{CO}_2\) measurements, that is, measurements from tall towers and by aircraft. Such an integrated effort has been undertaken within the Globalview-co2 project (Globalview-co2, 2009) by the Carbon Cycle Greenhouse Gases Group of the National Oceanic and Atmospheric Administration's Earth System Research Laboratory (NOAA ESRL). In addition to collecting raw data, the Globalview project extends their global product by applying a sophisticated interpolation scheme to fill temporal gaps in the station records (Masarie and Tans, 1995). Whenever I use atmospheric \(\text{CO}_2\) observations in this thesis, they are based on the Globalview-co2 database\(^1\). See figure 4.2 for an example of such \(\text{CO}_2\) networks.

By analyzing the air from tiny bubbles in ice cores, researchers have been able to reconstruct atmospheric \(\text{CO}_2\) concentrations from before the onset of direct instrumental measurements of air samples. The reconstructions reach as far back in time as 400,000 years and exhibit great precision (Petit et al., 1999). One of the major findings from reconstructed \(\text{CO}_2\) records is that atmospheric \(\text{CO}_2\) has varied quasi-periodically over the last 400,000 years between approximately 180 ppm and 280 ppm. These variations can be attributed clearly to the glacial-interglacial cycles, with the lower \(\text{CO}_2\) bound of 180 ppm corresponding to fully glacial conditions, and the upper bound of 280 ppm corresponding to interglacial conditions. This upper bound appears to be rather stable throughout interglacial times, with the current interglacial - the Holocene - being no exception. I show results from Antarctic ice core reconstructions for the most recent part of the Holocene, i.e. the last 1,000 years, in figure 1.1. The stable concentration around 280 ppm can be verified easily. Only after the onset

\(^{1}\)For more details on the Globalview-co2 database and data processing, refer to http://www.esrl.noaa.gov/gmd/ccgg/globalview/co2/co2_intro.html
Chapter 1. Introduction

of the Anthropocene in the mid-18th century the atmospheric CO$_2$ concentration began to rise exponentially, reaching 360 ppm already in the year 2000. The growth rate averaged 1.9 ppm yr$^{-1}$ over 2000-2008 (Le Quere et al., 2009) with substantial interannual variability (see also figure 4.1). In 2009, the concentration reached 387 ppm, an increase of more than 38% above the preindustrial level. The increase beyond preindustrial levels thus exceeds the 100 ppm variations between glacial and interglacial time periods, giving additional justification to coin the term "Anthropocene".

There are several reasons for why the observed increase in atmospheric CO$_2$ concentration is attributable to anthropogenic activities, in particular fossil fuel burning. First, the known sources of carbon are more than adequate to explain the observed increase. Actually additional carbon sinks are required to balance the global carbon budget, rather than other, unexplained sources of carbon (see section 1.1.2). Second, for thousands of years preceding the onset of the industrial era (but still within the current interglacial), the concentration of CO$_2$ in the atmosphere varied only by roughly ±20 ppm about the preindustrial 280 ppm value (Houghton, 2007). The timing of the rapid increase in the Anthropocene coincides with the rising carbon emissions from fossil fuel burning and land-use change. Third, the latitudinal gradient in CO$_2$ shows highest concentrations at northern mid-latitudes and lower concentrations further north and towards lower latitudes, consistent with the fact that most anthropogenic emissions are located in northern mid-latitudes (e.g. North America, Europe). This argument is further supported by the fact that this gradient has increased in proportion to emissions from fossil fuels (Keeling et al., 2005). And finally, the increase of CO$_2$ in the atmosphere is consistent with observed distributions of carbon isotopes and other biogeochemical tracers. For example, the Anthropocene increase in atmospheric CO$_2$ was accompanied by a decrease in the $^{14}$C content of CO$_2$ (Suess effect), which would be expected if the CO$_2$ added to the system stems from fossil fuel combustion. This is because fossil carbon is depleted in $^{14}$C through radioactive decay, owing to its age.

Before I turn into discussing the effects of the anthropogenic CO$_2$ increase in the atmosphere on the carbon cycles on land and in the ocean (figure 1.3), I describe here briefly its effect on Earth’s climate through the strengthening of the atmospheric greenhouse effect, because the latter was (and still is) one of the main reasons for researchers to study the global carbon cycle in the first place.

When averaged over the entire planet and a full day, the amount of solar energy reaching the top of the atmosphere is 342 W m$^{-2}$ (IPCC, 2007a), see also figure 1.2. In order for the planet to be in thermal equilibrium, the same amount of energy must, of course, be radiated back into space. However, the whole “trick” of the greenhouse effect is that this energy is radiated back at longer wavelengths than the incoming solar radiation and that the Earth’s atmosphere interacts with this longwave radiation in such a way that it “forces” the planet’s surface to become warmer than without atmosphere. About 30% of the sunlight reaching
1.1. The global carbon cycle in the Anthropocene

the top of the atmosphere is instantly reflected back to space. About 2/3 of the reflectivity are caused by clouds and small particles in the atmosphere called "aerosols"; the remaining 1/3 is due to bright (high albedo) surface patches, such as snow, ice and deserts. Major volcanic eruptions are one example of how important the reflective term in the energy budget can become, as they eject large amounts of aerosols into great heights, high enough to prevent them from being washed out by rain. The aerosols can then increase the reflectivity for about a year, before they fall back into the troposphere and get precipitated to the surface. By this process, major volcanic eruptions can cause a drop in mean global surface temperature of about 0.5°C, lasting for months or even a year. The remaining 70% of solar energy are absorbed by the Earth’s surface and atmosphere. This amount of energy is then radiated back into the atmosphere as longwave radiation. If the Earth had no atmosphere and could be approximated as a black body, its surface would have a temperature of -19°C (IPCC, 2007a), thus much colder than the observed global mean temperature of 14°C. The reason for the elevated temperature is the presence of an atmosphere containing greenhouse gases, which partially absorb and re-emit the longwave radiation. This is known as the natural greenhouse effect. The most important greenhouse gases are water vapor and CO₂, whereas the two most abundant constituents of the atmosphere - nitrogen and oxygen
- have no such effect. Clouds also exert a greenhouse effect as they are mainly compositied of water vapor; however, this effect is offset by their reflectivity, such that clouds tend to have an overall cooling effect on climate. Locally one can experience their warming effect: cloudy nights are often warmer than clear nights due to the clouds’ back-radiation of longwave energy. Anthropogenic activities intensify the natural greenhouse effect through the release of greenhouse gases, mainly CO$_2$, but also methane (CH$_4$) and nitrous oxide (N$_2$O). As mentioned above, the amount of CO$_2$ in the atmosphere has already increased by more than 38% over preindustrial values due to emissions from fossil fuel combustion and LUC. Thus, humankind has dramatically altered the chemical composition of the global atmosphere with substantial implications for climate. The climate sensitivity of our planet depends not only on the amount of greenhouse gases in the atmosphere, but also on a number of complex feedbacks that make precise projections of global warming difficult. Nevertheless, there is a strong consensus within the scientific community that the intensified greenhouse effect will lead to global warming and that it probably already accounts for most of the observed 0.6 ($\pm 0.2$) °C warming over the last century (Sarmiento and Gruber, 2002). In the most recent Intergovernmental Panel on Climate Change’s (IPCC) report the CO$_2$-induced greenhouse effect is identified to be by far the most important driver of climate change among all other radiative forcings (IPCC, 2007b).

### 1.1.2 Global carbon budget

The CO$_2$ released annually from fossil fuel burning and other industrial processes has increased nearly exponentially over the Anthropocene (i.e. the last 250 years), with only a few temporary interruptions in the trend caused by global-scale events such as the World Wars, the oil crisis in the 1970s or the collapse of the former Soviet Union in the early 1990s. Together with land-use change emissions approximately 530 Pg C (1 Pg C = 1 Gt C = $10^{15}$ grams of carbon) have been released in total over this period of time (to the end of 2008; Raupach and Canadell (2010)). Such an amount of carbon would have led to an increase in atmospheric CO$_2$ concentration of almost 250 ppm, if all of it had remained in the atmosphere (1 Pg C corresponds to a CO$_2$ concentration of 0.47 ppm). Yet, the observed increase is approximately 107 ppm, less than half (43% on average) of what would be expected. The atmospheric growth rate is lower because the terrestrial biosphere (plants and soils) and the ocean are continuously taking up the remaining - significant - fraction of the total emissions, therefore acting as sinks for anthropogenic carbon. Much progress has been made to quantify the relative role of these two major natural carbon sinks at the global scale. It appears a robust estimate that the "missing" carbon is about equally divided between them, at least over the last three decades (Sarmiento and Gruber, 2002; Le Quere et al., 2009). Figure 1.3 outlines the global carbon cycle with its three main reservoirs in the atmosphere, land and ocean, referenced to the 1990-1999 decade. Numbers are adapted from IPCC. Compared
1.1. The global carbon cycle in the Anthropocene

to the global total, there is much less understanding of the sinks’ spatial distribution and trend as well as mechanisms. An understanding of the mechanisms is key to predicting and mitigating the future impact of anthropogenic CO$_2$; the spatiotemporal distribution of sinks and sources in the carbon cycle is essential to gain insight into these mechanisms and to understand its past and current state.

![Global carbon cycle diagram]

**Figure 1.3** Global carbon cycle showing carbon stocks and fluxes between the three main compartments: the atmosphere, the ocean and the land. Anthropogenic contributions are in red; natural fluxes in black. Numbers represent an approximation for the 1990s average. NPP means net primary production. Figure from Sarmiento and Gruber (2002), numbers updated according to IPCC (2007a).

As depicted in figure 1.3, the carbon cycle comprises three major compartments (atmosphere, land, ocean) along with exchange fluxes between them and additional emissions from anthropogenic activities. The budget governing atmospheric CO$_2$ concentrations is as follows:

$$\frac{dC_a}{dt} = F_{Foss} + F_{LUC} - F_{LandSink} - F_{OceanSink},$$  \hspace{1cm} (1.1)

where $C_a$ is the atmospheric CO$_2$ content in Pg C. This budget expresses the mass balance constraint on atmospheric CO$_2$: the observed increase in atmospheric CO$_2$ must be equal to total inflows minus total outflows, where all fluxes have units of mass per unit time, that is, Pg C yr$^{-1}$ throughout this thesis. In the present era, i.e. the Anthropocene, the net inflows of CO$_2$ to the atmosphere are almost entirely due to anthropogenic emissions from fossil fuels and other industrial processes (mainly cement production; denoted $F_{Foss}$) and emissions from land-use change (denoted $F_{LUC}$). Net inflows from natural processes are much smaller,
although natural gross fluxes are the largest in the system and make up the “backbone” of the carbon cycle. However, natural fluxes are directed both in and out of the atmosphere with near cancellation, so that the net fluxes are close to zero (black arrows in figure 1.3). The outflows from the atmosphere are because of land and ocean CO$_2$ sinks (denoted $F_{\text{LandSink}}$ and $F_{\text{OceanSink}}$, respectively). Before I turn into discussing the terms in the carbon budget and how they can be quantified (see following section 1.2), I give here a short overview of the mechanisms on land and in the ocean that cause these two reservoirs to take up carbon from the atmosphere. For a more detailed description of the land sink I refer to (IPCC, 2007a, chap.7); the oceanic sink is explained in depth by Sarmiento and Gruber (2006).

### 1.1.3 Mechanisms behind the oceanic carbon sink

The processes controlling the uptake of carbon by the world’s oceans can be grouped into the following four:

- the ocean’s carbon chemistry,
- the air-sea CO$_2$ exchange,
- the vertical mixing between surface and deep waters, and
- the oceanic biological pump.

In the long term, the amount of atmospheric CO$_2$ taken up by the world’s ocean is controlled by the partial pressure of CO$_2$ (i.e. pCO$_2$) in the oceans. The oceanic pCO$_2$ is, in turn, set by a complex interplay of chemistry, biology and circulation. The comparison of reservoir sizes of the ocean and atmosphere (figure 1.3) suggests that the oceans should take up about 98% of the anthropogenic carbon added to the atmosphere. However, ocean chemistry ruins this optimistic view.

Dissolved inorganic carbon (DIC) in the ocean exists in three forms: dissolved CO$_2$ as well as bicarbonate (HCO$_3^-$) and carbonate (CO$_3^{2-}$) anions. The three forms are not equally abundant: in chemical equilibrium 99% of the DIC exist as bicarbonate and carbonate anions, whereas only 1% exists as dissolved CO$_2$. This asymmetric equilibrium (controlled by the underlying chemical reaction rates that transform all three forms into each other) is responsible for the high solubility of CO$_2$ in the ocean. But it is also responsible for the fact that the ocean will take up only 80% to 85% of the anthropogenic CO$_2$ in the atmosphere (Houghton, 2007; Sarmiento and Gruber, 2006). Oceanic carbon chemistry is the reason for the ocean’s buffer capacity.

The carbon exchange across the air-sea interface is controlled by the surface waters’ pCO$_2$. The direction of the net exchange depends on the direction of the pCO$_2$ gradient between
1.1. The global carbon cycle in the Anthropocene

surface waters and the above layer of air. Its magnitude is determined by both the steepness of this gradient and the local characteristics of the ocean’s surface, such as the level of roughness and bubble formation induced by wind speed conditions. A more detailed description of a CO$_2$ gas exchange model is provided in section 2.4; see also Takahashi et al. (2009b) for more information about the collection of pCO$_2$ data. Given a lower pCO$_2$ in surface waters compared to the above layer of air, the subsequent oceanic uptake of CO$_2$ is rather rapid with time scales from a few weeks to months, depending on average wind speed and the depth of the surface water layer. For example, the characteristic time for a 40 m deep surface layer to reach equilibrium with the atmosphere by air-sea gas exchange under average wind conditions is about 6 months (Sarmiento and Gruber, 2006).

If ocean circulation and mixing would be sufficiently fast, the short time scale of air-sea CO$_2$ exchange would allow the oceans to remain in equilibrium with the atmosphere, because the e-folding time of the anthropogenic CO$_2$ perturbation is much longer, i.e. several decades. However, the slow process of advection drives the mixing of surface waters with the deeper ocean, i.e. deep water ventilation. With the help of so-called age tracers, such as radiocarbon ($^{14}$C), it is possible to determine a ventilation timescale for the deep ocean of several centuries (Sarmiento and Gruber, 2006). Given this long timescale for deep water ventilation, one comes to realize that the time required for an atmospheric CO$_2$ perturbation (such as the anthropogenic one) to reach equilibrium in an ocean-atmosphere system must be on the order of a thousand years. Hence, the mixing between surface and deep layers is the bottle-neck for the oceanic uptake of CO$_2$, and it has enabled the atmosphere to be out of equilibrium with the oceans during the Anthropocene.

Although the oceanic CO$_2$ sink is dominated in the long-term by chemistry, in the mid-term by deep ocean ventilation and in the short-term by air-sea gas exchange, ocean biology is also significant. The biological pump transfers organic matter produced by phytoplankton in the surface layer to intermediate and deep waters. The net effect of the sinking and decomposition of organic matter is to enrich the deeper waters in DIC and CO$_2$ relative to surface waters and thus to reduce the CO$_2$ concentration of the atmosphere. The process is estimated to keep the concentration of CO$_2$ in the atmosphere considerably lower than what it would be in its absence (Sarmiento, 1993).

Based on tracer methods that account for biology in a stoichiometric way and separate anthropogenic DIC from the natural background, and combined with general circulation models, the net air-sea CO$_2$ exchange can be estimated by inverse modeling. Such an ocean interior inversion is presented in chapter 2 and underlies also the joint ocean-atmosphere inversions in chapters 3 and 4.

Trends in the uptake of anthropogenic CO$_2$ by the ocean over the recent decades as well as future uptake are primarily controlled by the evolution of the excess CO$_2$ concentration in the
atmosphere, as this builds up the pCO$_2$ gradient across the air-sea interface, which in turn is the thermodynamic driving force for the CO$_2$ uptake. In the absence of major feedbacks the oceanic sink grows along with growing atmospheric CO$_2$ content (Sarmiento and Gruber, 2006; Gloor et al., 2003). When estimating decadal trends in oceanic CO$_2$ uptake in chapter 4, I make use of this dominant influence of atmospheric CO$_2$ increase. However, changes in ocean temperature, salinity, circulation and the biological pump have the potential to alter oceanic CO$_2$ uptake as well (Le Quere et al., 2009).

1.1.4 Mechanisms behind the land carbon sink

The near-zero net natural background fluxes of carbon between terrestrial ecosystems and the atmosphere are largely the result of biological processes: photosynthesis and respiration. They amount to approximately 120 Pg C yr$^{-1}$ in each direction, about half of which is respired by the plants themselves, leading to the NPP (Net Primary Production) and (heterotrophic) respiration fluxes of about 60 Pg C yr$^{-1}$ in figure 1.3 (black arrows). Year-to-year variations in these fluxes owing to climatic variations, including variable fire frequency and intensity, can reach 5 Pg C yr$^{-1}$ (Peylin et al., 2005). As can be seen from figures 1.4 and 4.1, the variability in the background fluxes drives substantial, though lower than 5 Pg C yr$^{-1}$, interannual variability in the net land-atmosphere carbon exchange. However, net fluxes do not vary about a zero mean like one would expect from the near-cancellation of natural fluxes. Instead, it appears that terrestrial ecosystems acted as a net source of carbon until about 1940. After 1940 they have been a small net sink (Houghton, 2007), before they turned into a substantial sink in the beginning of the 1990s. The 1990s increase in land uptake will be discussed in detail in chapter 4 and is mentioned here only for illustration. This historical pattern of net land flux is quite different from the pattern of flux attributable to land-use changes: the latter has generally increased over time, a pattern not reflected in the net fluxes after 1940. Therefore, there must exist mechanisms causing the net land sink that are different from the background processes and overcompensate carbon emissions due to land-use change. The anthropogenic perturbation of the earth system seems an obvious reason for their existence. In the following I illustrate some of the mechanisms that have been proposed to explain the observed sink.

The proposed mechanisms responsible for carbon sinks on land belong to either one of the following two categories:

- enhanced growth driven by physiological or metabolic factors that affect rates of photosynthesis, respiration, growth and decay
- regrowth from past disturbances, historical land-use, or management, which affect the mortality of forest stands, the age structure of forests and, hence, their rates of carbon
1.1. The global carbon cycle in the Anthropocene

Physiological or metabolic factors that enhance growth rates
Carbon is taken up from the atmosphere through photosynthesis and released through respiration and fire. Respiratory processes include the respiration of plants, animals and microbes (largely soil respiration). Ecosystems will act as either sinks or sources of carbon, if an imbalance exists between these processes. For example, an increase in photosynthetic productivity will lead to an increase in carbon storage; at least until the carbon lost from the detritus pool comes into a new equilibrium with the intensified productivity. The time scale on which such an imbalance can sustain thus depends on the turnover time of carbon in an ecosystem: the longer the turnover time, the greater the disequilibrium and potential increase in storage.

CO$_2$ fertilization - induced by anthropogenic excess CO$_2$ in the atmosphere - is one potential mechanism to cause net carbon uptake by the land. Experiments have shown that many plants (trees, many crops, and vegetation in cold regions) exhibit increased rates of photosynthesis when exposed to elevated CO$_2$ (Houghton, 2007). In many cases this goes along with increases in productivity and biomass, although this is not universally observed. For example, Korner et al. (2005) found an immediate and sustained enhancement of carbon flux through 35-meter-tall temperate forest trees when exposed to elevated CO$_2$, but there was no overall stimulation in stem growth and leaf litter production after 4 years. Although growing vigorously, they found the trees not gaining more biomass carbon in stems. The pools of carbon in litter and soil also grow in response to elevated CO$_2$ (Luo et al., 2006).

Nitrogen (N) fertilization is caused by increased availability of biologically active forms of nitrogen (such as nitrate and ammonium) through anthropogenic activities, including the production of fertilizers. In regions where N availability is thought to be the limiting factor for NPP, this extra N is expected to increase NPP and also terrestrial carbon storage (Schimel et al., 1996). Examples for such regions are the temperate-zone ecosystems.

Climate change and variability lead to warmer temperatures and significant changes in precipitation patterns over large parts of the terrestrial biosphere. Warmer temperatures in concert with increased soil moisture often favor the growth of trees. In the long term these conditions also support the expansion of trees into tundra, savannas and grasslands. In cold ecosystems, such as tundra and taiga, an increase in temperatures can enhance productivity and carbon storage as well as lengthen the growing season (IPCC, 2007a).

The mechanisms described above often do not interact linearly in terms of changes in productivity or biomass storage, that is, they exhibit a range of couplings, both positive and negative. For example, CO$_2$ and nitrogen fertilization together have a greater effect on forest growth than the sum of their individual effects (Oren et al., 2001). An additional finding was
that the growth stimulation was relatively greater in nutritionally poor sites, suggesting that the indirect effects of the fertilization in stress attenuation may be more important than its direct effects on photosynthesis. An example of a negative coupling is provided by Shaw et al. (2002), who treated a California grassland with increases in temperature, precipitation, nitrogen deposition and atmospheric CO$_2$ concentration. While the increase of each individual factor led to enhanced NPP, their combined application led to a relatively reduced NPP enhancement. This can be interpreted such that elevated CO$_2$ increases productivity under poor growing conditions (i.e. colder temperatures, less precipitation, less nitrogen), but reduces it under favorable growing conditions. The most likely explanation given by Shaw et al. (2002) is that some soil nutrient became limiting, potentially due to increased microbial uptake. An important conclusion from these mechanisms and their couplings is that it is difficult, and often impossible, to attribute carbon sinks on land to individual, or combinations of, environmental factors influencing physiological or metabolic processes.

The role of ecosystem history and disturbances
Carbon sinks on land also result from the recovery of ecosystems from past disturbances. In addition to physiological and metabolic factors, the processes responsible for regrowth include succession, growth, mortality and aging. After a disturbance a phase of regrowth is initiated, with forests accumulating carbon as they grow. Changes in climate affect terrestrial carbon storage not only through physiological and metabolic effects on plant growth and respiration, but also through effects on stand composition and age structure in response to mortality and recovery from droughts, storms or fires. For example, an estimate of carbon uptake in the north American region (mainly USA), provided by Pacala et al. (2001), shows that the largest sink is due to regrowth in abandoned farmland and areas that had previously been logged. After regrowth in farmlands the second most important sink mechanism is the spreading of woody vegetation into areas where the frequency of fires has been reduced (or suppressed).

Over the recent years the most common explanations for the net land carbon sink within the scientific community have shifted from processes affecting the physiology of plants and microbes towards land-use practices and disturbances that influence the age structure of ecosystems. One reason was that CO$_2$ fertilization appears to be less important in forests than in short-term greenhouse experiments (Oren et al., 2001). Another reason was that ecosystem models based on plant physiology and driven by elevated CO$_2$ and changes in climate were able to explain only a small fraction of the observed carbon accumulation in north American forests (Schimel et al., 2000; Pacala et al., 2001), with a much greater influence of past changes in land-use. In conclusion, the physiological effects of CO$_2$, N and climate have either been unimportant or their effects have been offset by other, unknown, influences.
1.2 Net carbon fluxes between atmosphere, land and ocean

In this section I discuss the terms in the global budget equation (1.1) and how they have been quantified in scientific studies, with special focus on the ocean and land sinks. A recent synthesis by Le Quere et al. (2009) shows each term for the period 1850 to 2008, estimated based on a mixture of bottom-up methods and numerical modeling. By convention, a sink of CO$_2$ is reported as negative (removing CO$_2$ from the atmosphere) and a source as positive (adding CO$_2$ to the atmosphere). An exception is the atmosphere itself: it is denoted as sink for CO$_2$ with regard to its accumulation of CO$_2$. The accuracy to which each sink/source term can be estimated varies strongly among the individual terms. The atmospheric CO$_2$ growth rate and global fossil fuel emissions are the best known terms in the carbon budget, followed by estimates of land-use change and the oceanic carbon sink. The carbon sink over land is by far the least well-known component (Sarmiento et al., 2010).

![Major fluxes in the global carbon cycle over (roughly) the Anthropocene. Each flux contribution corresponds to a term in the global CO$_2$ budget (1.1). Anthropogenic CO$_2$ emissions are shown as positive fluxes into the atmosphere; they comprise emissions from fossil fuel combustion and other industrial processes (such as cement manufacture and gas flaring; $F_{\text{Foss}}$) and emissions from land-use change (mainly deforestation and after the mid-1980s mainly tropical deforestation; $F_{\text{LUC}}$). Negative fluxes denote the accumulation of carbon in the three major compartments: the atmosphere ($dC_a/dt$), the land biosphere ($F_{\text{LandSink}}$) and the ocean ($F_{\text{OceanSink}}$). Numbers in the right panel give the 2000-2008 average estimates from Le Quere et al. (2009). Figure of Raupach and Canadell (2010), based on Le Quere et al. (2009).](image-url)
1.2.1 Sources of carbon

The increase in atmospheric CO$_2$ concentration (or atmospheric growth rate; $dC_a/dt$) is the best known component of the carbon cycle. The fact that CO$_2$ is a long-lived, passive tracer in the atmosphere, together with rapid atmospheric circulation, enables researchers to obtain global CO$_2$ growth rates from a few well-situated measurement sites (cf. 1.1.1). The uncertainty of the atmospheric CO$_2$ growth rate is inherited from the measurement error of direct observations of CO$_2$ concentration, which is only a few percent. In chapters 3 and 4 I average CO$_2$ records from two stations, Mauna Loa and South Pole, to obtain global growth rates. This adds some uncertainty due to the inter-station spread, however, overall uncertainty is still smaller than ±10% on an annual basis (see background patches in figure 4.1), and decreases further when averaged over a decade. As can be seen from both figures 4.1 and 1.4, interannual variability in the atmospheric CO$_2$ content is substantial, owing to the variable land sink. As discussed in the previous section the main driver of these variations is climate variability, which in turn is dominated by the El Nino/Southern Oscillation or ENSO (Peylin et al., 2005; Baker et al., 2006).

Carbon emissions from fossil fuel combustion, cement production and gas flaring ($F_{\text{Foss}}$) are also a well-constrained component of the carbon budget. In figure 1.5 I show the global distribution of $F_{\text{Foss}}$ for the year 1995. From the zonally integrated emissions in the right panel the dominant role of the industrial countries in burning fossil fuels becomes apparent. Fossil fuel burning increased at a rate of approximately 1.0% per year during the 1980s and 1990s, before it accelerated to about 3.8% per year over the period 2000-2008 (Boden et al., 2010). The annual record of fossil fuel emissions is shown in figure 4.1 throughout 1979-2008. The uncertainty of global fossil fuel emissions is considered to be about ±6% (Marland et al., 2008), translating into approximately 0.5 Pg C yr$^{-1}$ for an 2000-08 emission estimate of 7.7 Pg C yr$^{-1}$.

Emissions from LUC ($F_{\text{LUC}}$) are the second-largest anthropogenic source of CO$_2$. They result from deforestation, logging and intensive cultivation of cropland soils, partly compensated by CO$_2$ uptake from the regrowth of secondary vegetation as well as the rebuilding of carbon pools following afforestation, abandonment of agriculture and fire suppression. LUC emissions for a particular year reflect not only deforestation rates during that year, but also carry-over effects of CO$_2$ losses from areas deforested in previous years. That is a major difference to fossil fuel emissions reflecting only instantaneous economic activity, and is the main reason why LUC emission estimates are less certain than fossil fuel emission estimates. In their recent synthesis, Le Quere et al. (2009) combine LUC-induced net CO$_2$ fluxes based on United Nations data with a book-keeping method$^2$ (Canadell et al., 2007; Houghton, 2003) to estimate net LUC CO$_2$ emissions to 1.5 (±0.7) Pg C yr$^{-1}$ for the period

$^2$Book-keeping methods track the amount of carbon released to the atmosphere from clearing and decay of plant material, and additionally account for the amount of carbon accumulating as vegetation regrows.
1.2. Net carbon fluxes

Figure 1.5 Global distribution of fossil fuel emissions for the year 1995. The color scale is logarithmic. The right panel shows the zonally integrated fossil fuel emissions. The emerging pattern demonstrates the dominant fossil fuel sources from industrial countries in temperated northern hemisphere regions. The figure is based on data from Brenkert et al. (1998).

1990-2005. The dominant driver for these emissions was tropical deforestation (about 1.3 Pg C yr\(^{-1}\)), which consists of two components: the conversion of tropical forests to pastures or croplands and timber exploitation. The overall LUC CO\(_2\) source (tropical and non-tropical) remained relatively constant over the last three decades, though its contribution to total anthropogenic CO\(_2\) emissions decreased from 20% in the 1990s to 12% in 2008 owing to the strong increase in fossil fuel emissions. In 2008, total anthropogenic CO\(_2\) emissions reached 10 Pg C yr\(^{-1}\) (more precisely: 9.9 (±0.9) Pg C yr\(^{-1}\)).

1.2.2 Sinks for carbon

The estimation of carbon sinks on land (\(F_{\text{LandSink}}\)) and in the oceans (\(F_{\text{OceanSink}}\)) is a large scientific field with lots of studies conducted based on a variety of methods, owing in particular to the key role of CO\(_2\) surface fluxes (land and ocean) in the climate system. The sum of the ocean and land sink can be obtained by subtraction of the atmospheric CO\(_2\) growth rate from the total CO\(_2\) release to the atmosphere, due to the relatively small uncertainties associated with the source terms. During the Anthropocene the global carbon sink accounted for a little more than half of the sources (cf. section 1.1.2), but what can we say about the individual contributions from the land and ocean as well as their regional distribution and
temporal variations?

Here I present an overview on the fundament of work to quantify these sinks individually, before I will profile my own approach in the following section 1.3. I describe first how the global carbon sink can be separated into a terrestrial and an oceanic contribution, before I turn more into their spatial distribution together with temporal changes on timescales of months to decades. More detailed discussions in the context of my own work are provided in the discussion sections of chapters 3 and 4.

Methods to quantify CO$_2$ exchange between the earth’s surface and the atmosphere, based on direct observation, can generally be categorized as

- bottom-up (based on measurements in the land or ocean compartments), or
- top-down (based on measurements in the atmosphere).

The use of biogeochemical or ecosystem models is an alternative approach to estimate CO$_2$ exchange. Space-based remote sensing of carbon-related parameters, such as forest greenness or ocean color, cannot strictly be considered bottom-up, because models are needed in addition to translate the color data into meaningful parameters that can be used to derive carbon fluxes. For example, the leaf area index (LAI) may be deduced from the greenness data, then translated into growth rates of the vegetation (see Myneni et al. (2007) for an example). Similarly, ocean color data can be used to estimate phytoplankton abundance, which must then be fed into an ocean biogeochemistry model to estimate export production and the subsequent CO$_2$ drawdown from the atmosphere (for example data from SeaWiFS: http://oceancolor.gsfc.nasa.gov/SeaWiFS).

The atmosphere “sees” both the ocean and land compartments, so should it not be possible to separate both sinks from each other by just measuring the CO$_2$ in the atmospheric region directly above a chosen land or ocean region? Unfortunately, due to the short timescale of atmospheric circulation, the regional CO$_2$ drawdown gets mixed around the globe so quickly that its origin cannot be localized easily. However, if the circulation is taken into account, the origin of the CO$_2$ exchange can be retraced; this is the basic principle of an atmospheric inversion. On very small spatial scales it is also possible to measure the CO$_2$ exchange directly, for example using flux towers.

At global scale it is possible to use atmospheric concentrations of oxygen (O$_2$) to separate the ocean from the land sink without the need to apply circulation models, owing to the different uptake processes on land and across the air-sea interface (IPCC, 2007a). The exchange of carbon and oxygen between the land biosphere and the atmosphere is inversely coupled, due to the O$_2$:C stochiometry that underlies the processes of photosynthesis and respiration. By contrast, the oceanic uptake of carbon is independent of the uptake of oxygen, because
it is driven by the physical process of gas exchange. The global land sink can be inferred directly from the biospheric imprint on atmospheric O$_2$. Based on parallel measurements of atmospheric CO$_2$ and O$_2$ (usually in the form of the O$_2$/N$_2$ ratio), studies show that in the 1990s, the land must have taken up less carbon than the ocean, in particular when a possible climate-induced net outgassing of oceanic oxygen is taken into account (Plattner et al., 2002; Bopp et al., 2002). The work by Manning and Keeling (2006) indicates that atmospheric O$_2$ is decreasing at a faster rate than CO$_2$ is increasing, which demonstrates the importance of the oceanic carbon sink.

Another method for determining the partitioning of anthropogenic CO$_2$ between ocean and land is based on measurements of carbon isotopes, specifically the ratio of $^{13}$C/$^{12}$C, in the atmosphere and ocean. Fossil fuel CO$_2$ is slightly depleted in $^{13}$C relative to $^{12}$C, leading to a continuous decrease of the $^{13}$C/$^{12}$C isotopic ratio of atmospheric CO$_2$. The CO$_2$ with this isotopic signature is then redistributed within the global carbon cycle, permitting us to distinguish anthropogenic CO$_2$ from the natural CO$_2$. The carbon uptake across the air-sea interface does not fractionate, that is, the gas exchange process does not distinguish $^{13}$C from $^{12}$C. By contrast, land uptake fractionates, because land plants prefer taking up the lighter isotope $^{12}$C over the heavier $^{13}$C, thereby increasing the atmospheric $^{13}$C/$^{12}$C isotopic ratio. If knowledge about this fractionation is combined with accurate estimates of fossil fuel CO$_2$ emissions, measurements of the $^{13}$C/$^{12}$C ratio in the atmosphere allow to separate the ocean from the land sink (Battle et al., 2000).

At smaller spatial scales bottom-up approaches to estimate the land sink include eddy covariance measurements from flux towers (Hutyra et al., 2007) and carbon inventory studies (Phillips et al., 2009; Lewis et al., 2009; Ciais et al., 2010). The eddy covariance technique is used to measure and calculate vertical turbulent fluxes within atmospheric boundary layers. It is a statistical method that analyzes high-frequency wind and scalar atmospheric CO$_2$ data series, and yields values for vertical CO$_2$ fluxes. It relies on two basic assumptions: 1) the mean vertical flow is negligible over horizontal homogeneous terrain (vertical transport only via turbulence) and 2) air density fluctuations are negligible. The technique allows for local measurements of the net vertical CO$_2$ exchange between the land surface and the atmosphere. The tower height determines the size of the flux footprint, that is, the surface area for which the calculated fluxes are representative. Both the eddy covariance and the inventory methods are very valuable to accurately estimate local to regional carbon budgets, but they represent only local vegetation type and are, hence, difficult to upscale to continental or global regions characterized by heterogeneous biome structure.

Bottom-up estimates of terrestrial carbon sinks are only available since about 1990 (Sarmiento et al., 2010). The most complete inventories exist for northern mid-latitudes. Forest inventories in northern mid-latitudinal lands systematically measure wood volumes from overall more than a million plots (Houghton, 2007). A synthesis study of these inventories attributed
Chapter 1. Introduction

A net sink of -0.6 to -0.7 Pg C yr\(^{-1}\) to northern mid-latitude forests for the years around 1990 (Goodale et al., 2002). The North American sink around the year 2003 was estimated to -0.5 Pg C yr\(^{-1}\) by Pacala et al. (2007). They attributed about 60% of this sink to the forest sector. This forest sink was primarily due to increases in the mean age of forest stands, while fertilization by N or CO\(_2\) showed no clear effect. The two studies agree on the sink estimate for North American forests, though a direct comparison should be done with caution owing to the different time periods of study. Sarmiento et al. (2010) augmented these results with estimates for the non-forest sector of Eurasia and fluxes from boreal regions of both continents. They arrived at an estimate of -1.0 Pg C yr\(^{-1}\) for the combined northern temperate and boreal zones, with uncertainty of at least 50%. The inventory-based estimates have been confirmed by eddy-covariance studies (Pacala et al., 2007), yet this confirmation is only possible locally and cannot be expanded to check whether it is possible to come up with consistent continental-scale bottom-up sinks from both inventory and flux tower studies. The main reasons are the limited spatial extension of inventories and the limited size of the flux tower network, preventing accurate spatial upscaling. In addition, the eddy covariance technique fails during calm nights with little turbulence and predominantly lateral flow, thus violating the basic assumptions of the technique (Saleska et al., 2003).

Inventory data are available for tropical forests as well, though to much less extent than for the northern temperate zones. Based on a network of 60,000 trees across the Amazonian rainforest, Phillips et al. (2009) estimate that undisturbed forests took up carbon during the 1990s at an average rate of -0.8 Pg C yr\(^{-1}\). By augmenting this estimate with similar data for tropical regions in Asia and Africa (Chave et al., 2008), Sarmiento et al. (2010) extrapolate an overall sink in mature tropical forests of about -1.4 Pg C yr\(^{-1}\). When combined with LUC emissions due to the pantropical deforestation of 1.3 Pg C yr\(^{-1}\) (see previous section, Le Quere et al. (2009)), this gives a near-neutral net carbon balance for tropical land regions. Whether the non-LUC sink is due to physiological factors of enhanced growth (such as CO\(_2\) fertilization), or due to a recovery from a large scale disturbance event, is debated in the literature (Malhi, 2010). The near-neutral net tropical balance is in broad agreement with a 4 year flux tower study from Amazonia in the early 2000s (Hutyra et al., 2007). Yet again, a direct comparison is difficult due to the non-overlapping time periods of the studies. And still the above mentioned limitations of land bottom-up methods remain. It is particularly difficult to estimate decadal trends in the regional land carbon budget based on inventory or flux tower studies, because of these inconsistent study periods and because of the young history of these approaches.

One alternative to these bottom-up methods are biosphere models, which can be used to simulate present day as well as trends in the land carbon sink (Potter et al., 2003; Le Quere et al., 2009; Piao et al., 2009). However, the underlying drivers such as climate, CO\(_2\) and land-use change, are afflicted with uncertainties that lead to a considerable spread of results.
1.2. Net carbon fluxes

The air-sea flux of anthropogenic CO\textsubscript{2} has been estimated using forward simulations of general ocean circulation models coupled to biogeochemistry models and forced by the observed atmospheric CO\textsubscript{2} perturbation. An example are simulation results from 13 ocean biogeochemistry models that participated in the 2nd phase of the Ocean Carbon-cycle Model Intercomparison Project (OCMIP-2) (Watson and Orr, 2003). It is in particular the inclusion of the ocean biogeochemistry model that introduces considerable uncertainty to the simulated net air-sea CO\textsubscript{2} exchange, because it has to embrace all the processes described in section 1.1.3. By contrast, the physics underlying the circulation models are well-known. The main limitations for circulation models include insufficient resolution and uncertain surface forcings. One idea was, therefore, to use passive carbon tracers in the ocean interior that do not react chemically and are not altered by biology to infer air-sea CO\textsubscript{2} fluxes. In chapter 2 I describe one of these approaches: an ocean interior inversion of the so-called C\textsuperscript{*} tracer developed by Gruber et al. (1996). This ocean interior inversion is a key component of the inverse models I developed in the framework of this thesis, which is why I explain the method in a separate chapter.

Among the bottom-up methods in the ocean are carbon inventories. Using inorganic carbon measurements from a global network of cruises together with the tracer-based method to separate natural from anthropogenic carbon of Gruber et al. (1996), Sabine et al. (2004) estimate the global longterm (averaged over the period 1800-1994) oceanic sink of anthropogenic CO\textsubscript{2} to -0.6 Pg C yr\textsuperscript{-1}. It is generally difficult to obtain inventory-based oceanic flux estimates on shorter timescales, due to the sluggish ocean circulation and the huge background carbon stock. A more direct and often used approach, albeit without ability to separate the net flux into its natural and anthropogenic components, is the measurement of the air-sea difference of pCO\textsubscript{2}, which when combined with bulk gas exchange parameterizations yields an estimate of the net flux (see section 2.4).

Another alternative to estimate the regional distribution of carbon sources and sinks is the top-down atmospheric inversion method, where a set of regionally resolved ocean-atmosphere and land-atmosphere carbon fluxes are adjusted to be optimally consistent with the observed atmospheric CO\textsubscript{2} distribution. Atmospheric circulation models are needed to link surface fluxes to concentrations in the atmosphere. However, these models are purely physical without need of a biogeochemical component, since CO\textsubscript{2} is a passive tracer in the atmosphere. Given a perfect circulation model together with an unlimited and perfectly distributed set of atmospheric CO\textsubscript{2} observations, such an atmospheric inversion could provide CO\textsubscript{2} flux estimates at any desired spatial and temporal scale. However, both the transport models and the observational networks are far from being perfect, posing limitations on this method. Here I will mention a few previous atmospheric inversion studies. In chapter 2 I will then explain its principles in more detail, because an atmospheric CO\textsubscript{2} inversion underlies all of my own inverse models in chapters 3 and 4.
Atmospheric inversions interpret gradients in atmospheric CO$_2$ concentration. The most pronounced gradient exists in north-south direction, because CO$_2$ from fossil fuel burning is released primarily in the northern mid-latitudes (figure 1.5), thus inducing a permanent accumulation of CO$_2$ in the Northern Hemisphere. In the first inverse studies, Keeling (1960) and Tans et al. (1990) detected that the observed latitudinal gradient of about 2 ppm was smaller than expected from fossil fuel emissions alone (3.8 to 5.6 ppm). Both authors thus concluded that, in order to match the measured CO$_2$ gradient, a sink was needed north of the equator. However, they did not agree on whether this sink is located over land or ocean. In the early 1990s, uncertainties were larger for bottom-up estimates than for atmospheric estimates. Only very sparse measurements of land fluxes (and of ocean fluxes) were available to confirm the results deduced from the atmospheric CO$_2$ distribution. During the 1990s, more and more atmospheric stations were deployed over the Northern Hemisphere and enabled atmospheric inversions to improve estimates of CO$_2$ sinks and sources (Bousquet et al., 1999b; Gurney et al., 2002, 2004; Roedenbeck et al., 2003a; Baker et al., 2006; Peters et al., 2007) both with regard to spatial scales (continental scale instead of broad latitudinal bands) and temporal resolution (monthly or even weekly instead of longterm annual mean). The results generally confirmed the existence of a large Northern Hemisphere sink of about -2.5 Pg C yr$^{-1}$ and showed that a significant fraction of that sink is in terrestrial ecosystems, not in the ocean. The uncertainty is large, about 2 Pg C yr$^{-1}$, due to the still limited number of atmospheric stations and to systematic differences among transport models.

Indeed, the density of atmospheric CO$_2$ data stations is key to any atmospheric inversion. The spatiotemporal resolution accessible with an inversion directly depends on the density (in space and time) of the data. In addition, any successful inversion requires an accurate description of atmospheric transport, including the vertical mixing. At present, atmospheric inversions have been used to estimate the CO$_2$ surface fluxes at continental scales from an observational network of about 100 stations (for example the Transcom series of inversions, as described below). The growing interest in carbon fluxes in the context of climate change has led to a rapid expansion of the network. Furthermore, the temporal measurement density has begun to increase from typically weekly or biweekly intervals to continuous sampling. Yet, there are two persisting problems with the current network development. First, dense and continuous networks exist primarily over two regions only: Europe and North America. Second, the network will never have the density required for global monitoring at fine scales of, say, 100 km. In particular, the in situ monitoring network will not be sufficiently expandable over the oceans and over large forest areas that are difficult to access, such as the Amazon, Africa or Siberia. It is therefore appealing to consider using spaceborne observations as well, because they provide a high density of measurements over large parts of the globe. In addition, they are free from local pressure (including political issues) and can provide uniform coverage over all regions. However, satellite-derived CO$_2$ measurements often exhibit low accuracy, sometimes so low that they do not add any new information about CO$_2$ surface
fluxes compared to surface station inversions.

One of the most prominent top-down studies is the Transcom-3 inversion intercomparison, which estimated carbon sinks and sources for 11 oceanic and 11 land regions globally and averaged over the 1992-1996 period (Gurney et al., 2003, 2004). Choosing this period was guided by the wish to avoid major disturbances of the carbon cycle: it starts one year after the Pinatubo eruption in 1991 and ends one year before the major El Nino event in 1997/8. The Transcom-3 inversions generally find a strong net carbon source in the tropical land and a large net sink in the northern hemisphere extra-tropics, contrasting the bottom-up estimates. Within the respective uncertainty bands the Transcom and bottom-up estimates overlap, but there is a rather large offset of about 1 Pg C yr\(^{-1}\) between them with regard to both the tropics and the northern land regions.

Apart from these discrepancies between top-down and bottom-up estimates for air-land CO\(_2\) fluxes, large discrepancies also exist for air-sea CO\(_2\) fluxes. The Transcom estimates for air-sea fluxes disagree with independent estimates obtained by air-sea pCO\(_2\) difference measurements combined with a gas exchange model (Takahashi et al. (2009b); see section 2.4 for more information on the method). They also disagree with estimates from ocean interior DIC inversions (Gruber et al. (2009); see chapter 2 for more information on the method). The disagreement is particularly striking in the tropical oceans, where the Transcom models underestimate the CO\(_2\) release from the water relative to the bottom-up pCO\(_2\) and ocean inverse estimates. Other examples are that the Transcom models tend to underestimate oceanic uptake in the Southern Hemisphere temperate basins, and tend to overestimate the oceanic uptake in the Southern Ocean (Gruber et al., 2009).

### 1.3 Profiling this Ph.D. project

My thesis is motivated by the existing discrepancies between top-down and bottom-up methods for the quantification of CO\(_2\) fluxes between the atmosphere and both the land and the ocean. Together with remote sensing methods they belong to the data-based approaches, although models are involved. But the models are not the dominant part of the methods, contrasting process-based model approaches for the land ecosystems or the biogeochemical cycles in the ocean.

Process-based models can be used to estimate carbon exchange on all spatiotemporal scales, but are afflicted with uncertainties (e.g. process modelization or applied forcings) and always need careful validation with observation-based methods. Models are ideal to study processes that lead to CO\(_2\) fluxes and to attribute flux changes to individual processes, but data-based methods are preferable to estimate net fluxes (given that enough data are
available). Recently, Le Quere et al. (2009) achieved a closure of the global carbon budget by using only process-based models to estimate the land and ocean sinks. Their residual flux, that is, the portion of CO$_2$ flux that cannot be explained by the models, is remarkably close to zero when averaged over the last 50 years. However, on an annual and decadal basis the residual flux is significant, leaving up to 1 Pg C yr$^{-1}$ unexplained.

In my work process-based models do not play an integral role, although results are compared to them in the discussion sections of chapters 3 and 4. Instead, the main focus is to try to reconcile air-sea and air-land CO$_2$ exchange from data-based methods by incorporation of several data streams into a joint ocean-atmosphere inversion system. In addition to the multi-data stream approach, several atmospheric and oceanic circulation models are compared to assess transport error. In chapter 4 I put also particular emphasis on assessing the sensitivity of flux results to the choice of atmospheric CO$_2$ observation sites. In summary, my Ph.D. project distinguishes itself from previous work by the incorporation of all of the following features:

- Joint ocean-atmosphere inversion that interprets observational data from both the atmosphere (atmospheric CO$_2$ measurements) and the ocean (ocean interior DIC and surface ocean pCO$_2$ measurements).
- Additional constraint from bottom-up land uptake data.
- Monthly flux resolution averaged over 5 to 10 years (cyclostationary).
- Estimation of decadal flux trends over the last 3 decades.
- Multi-model approach: 11 to 12 atmospheric transport models are used in combination with 10 oceanic transport models, thus enabling assessment of transport error.
- Assessment of impact of each data constraint.
- Assessment of inversion sensitivity to the composition of the atmospheric CO$_2$ station network.

The aforementioned discrepancy in air-sea net CO$_2$ fluxes between the Transcom atmospheric inversion and bottom-up fluxes obtained from air-sea differences in pCO$_2$ and inversion of ocean interior DIC is in part due to the fact that the Transcom inversions are generally able to provide well-constrained fluxes for latitudinal bands, but have difficulties to attribute these fluxes to oceanic or land parts within the bands. The main reason is the fast atmospheric mixing rates within such bands, thus diluting flux contributions from ocean and land and making it difficult for the inversion to distinguish between the two. On the other hand, estimates from an ocean inversion put strong constraints on basin-scale air-sea fluxes. Therefore, by combining the two approaches, the flux contribution from the land parts within
1.3. Profiling this Ph.D. project

a latitudinal band can be obtained by subtracting the air-sea flux from the overall flux inside the band. This also makes the land fluxes consistent with measurements taken in the ocean interior. Jacobson et al. (2007a) were the first to combine an ocean with an atmosphere inversion to obtain air-sea and air-land fluxes that were simultaneously consistent with carbon measurements from both atmosphere and ocean. However, their approach remained limited to annual mean fluxes. I overcome this limitation by extending the method to the seasonal timescale with the help of monthly atmospheric CO$_2$ data and the monthly pCO$_2$ climatology of Takahashi et al. (2009b).

Studies aiming at seasonal or higher temporal resolution have so far either relied upon one data stream (e.g. Gurney et al. (2004); Baker et al. (2006); atmosphere-only inversion) or put the emphasis on confined areas rather than global flux estimates (e.g. Peters et al. (2007); North America). Using monthly data and resolving monthly or seasonal fluxes is desirable because of two reasons. First, it has been shown that a substantial part of transport uncertainty in annual inversions can be attributed to seasonally unresolved atmospheric CO$_2$ concentrations and winds (Gurney et al., 2003). Second, annual atmospheric inversions tend to be biased compared to seasonal inversions because of the so-called seasonal rectifier effect, described by (Denning et al., 1995). This effect represents the co-variation between the seasonality of terrestrial CO$_2$ fluxes and atmospheric transport, in particular seasonal variations in the planetary boundary layer (PBL) depth and dynamics. A latitudinal CO$_2$ gradient is produced, even with purely seasonal, annually balanced, land biospheric fluxes. Because transport models in annual inversions do not resolve these PBL changes, the utilized annual CO$_2$ data are usually corrected for this effect by subtracting a CO$_2$ concentration field representing the rectifier effect, before inverting the data. However, this rectifier field is based on ecosystem model simulations and inherits all the model’s uncertainty.

The uncertainty in atmospheric transport is an important source of error in atmospheric inversions, though difficult to quantify directly. It is usually assessed by model intercomparison, as in the Transcom studies. I followed this approach, too, by incorporating 11 to 12 of the original Transcom-3 (level 2) models in the atmospheric part of the joint inversion, and 10 oceanic transport models from the Ocean Inversion Project$^3$ (OIP). In a recent study that made use of new vertical CO$_2$ profiles in the atmosphere (Stephens et al., 2007), three of the Transcom atmospheric transport models were found to be consistent with the annual mean observed vertical CO$_2$ gradients. Although their results are still debated, I exploit the changes in flux estimates induced by restricting the pool of atmospheric transport models to this subset (the so-called Stephens subset).

Studies aiming at decadal trends of CO$_2$ fluxes on a global scale are much less abundant in the scientific literature than estimates for one specific period. Le Quere et al. (2009) investigated the global carbon budget over the last five decades based on process models,

---

$^3$For more details on the project, visit http://quercus.igpp.ucla.edu/OceanInversion
but they made no attempt to distribute carbon sources and sinks spatially, nor did they resolve the seasonal variability. In a recent study Gurney and Eckels (2011) focused on the last three decades. Their inversion is an extension of the Transcom interannual atmospheric inversion (Baker et al., 2006) to the time period 1980-2008. It is monthly resolved, but purely atmospheric. In addition, it relies heavily on additional model-based estimates of terrestrial fluxes, which I was able to avoid by using the oceanic data streams. Only few bottom-up estimates exist that span several decades (owing to the young history of inventories and eddy flux studies), and those that exist cover selected regions only.

It has been shown that the selection of atmospheric CO$_2$ stations has a considerable impact on CO$_2$ fluxes from atmospheric inversions (Roedenbeck et al., 2003a; Gurney et al., 2008). Interannual variability as well as longterm trends turned out to be more robust against changes in the station network composition. Assessing this impact on my joint inversion estimates is also an important part of this Ph.D. thesis.

As outlined before, the consideration of satellite-derived measurements of atmospheric CO$_2$ concentration in an atmospheric inversion could improve our knowledge on CO$_2$ surface fluxes significantly, because of the potentially large density and homogeneous distribution of measurements. The sheer amount of data may outweigh the limited accuracy associated with remote CO$_2$ observations (Kadygrov et al., 2009; Breon and Ciais, 2010). Why is a high accuracy needed? The main reason is because CO$_2$ is a long-lived greenhouse gas in the atmosphere. Long-lived means that the time scale of atmospheric mixing is much shorter than the residence time of CO$_2$ in the atmosphere. This leads to a high background concentration compared to the concentration gradients induced by local sources and sinks of CO$_2$ at the surface. This is true in particular in higher atmospheric layers, where the surface imprint is very weak. Yet, these high layers are important for a satellite inversion, because spaceborne instruments measure column-mean CO$_2$ concentrations rather than those in the boundary layer. In order to resolve the small relative gradients a high accuracy is needed. As part of chapter 3 I will discuss an inversion based on pseudodata from an idealized satellite. A study based on real measurements was not possible since no high-quality product of satellite CO$_2$ observations exists to date (partly due to the regrettable crash of NASA’s OCO satellite in 2009). The main emphasis will be on estimating a measurement accuracy that would be needed in order for the satellite data to compete with the present surface network. A key role will play how the situation changes if the ocean constraint is taken into account in addition to the atmospheric CO$_2$ network.

I structured my Ph.D. thesis as follows: this introductory part is followed by a methodological chapter, where I describe the atmospheric inversion, the ocean inversion and how a joint estimate is constructed. In chapter 3 I present the results from the joint inversion for the 1992-1996 period, before I turn to decadal flux variability in chapter 4. Both chapters represent independent papers in pre-submission status; their submission is scheduled for right after
1.3. Profiling this Ph.D. project

the completion of my Ph.D. project. They are written in paper form, that is, with their own abstract, introduction, method section and conclusions. However, their method sections were shortened to only contain parts that are specific to each paper and not already included in chapter 2. In chapter 5 I close with a general summary and outlook.
Chapter 2

Inverse method for carbon exchange across Earth’s surface

2.1 Data assimilation and Bayesian concept

The atmosphere-only as well as the joint atmosphere-ocean inversions I developed over the course of my Ph.D. represent two techniques in the field of data assimilation. The atmosphere-only inversion solves for CO$_2$ surface fluxes using the least-squares method, while the joint inversion follows the Bayesian concept of combining a priori (henceforth, prior) flux information with CO$_2$ concentration data. Many additional assimilation techniques have been developed, in particular in the fields of meteorology and oceanography, which I do not include here. For the reader interested in more details about the basic statistical concepts I recommend the seminal textbook contributions of Tarantola (2005a) and Enting (2002). For a more complete overview on assimilation techniques and applications in atmospheric and oceanic systems I refer to Enting (2002) and Le Quere and Saltzman (2010).

Data assimilation is an analysis technique that combines observations distributed in space and time (for example atmospheric CO$_2$ concentrations) with a dynamic model. It aims at the approximation of the true state of a physical system at a given time. The true state would be known in a perfect world, where we had a perfect model that correctly reproduces perfect data. The model thus would generate a fully consistent four-dimensional picture of the real ocean or atmosphere (or any other physical system considered). However, in the real world we have to cope with imperfect models (mostly systematic deficiencies) in concert with imperfect data (e.g. measurement errors, methodological uncertainties) and often poor data coverage (in both space and time). Data assimilation is then used to combine data and model to produce an "optimal" estimate of the system’s state, where the optimum is usually defined by the minimization of penalty, or cost, based on metrics underlying the so-called cost function (denoted $J$ hereafter).
The concepts of data assimilation can be applied to different types of models, for example atmospheric models and ocean circulation models (as in the thesis presented here), wave models, sea surface temperature or gas exchange models, models of land surface properties, or vertical column models of the atmosphere for satellite data retrieval. Any model links input parameters (model state) to output parameters (observations) by

\[ d_{\text{mod}} = Q(x) \]  

(2.1)

where \( Q \) is the model operator. \( Q \) operates on the input vector \( x \) and generates the output vector \( d_{\text{mod}} \) of modeled data (observations). The input vector contains the set of variables that are needed to completely specify the model state (e.g. regional air-sea or air-land \( \text{CO}_2 \) fluxes). As for now, the model operator may be linear or nonlinear; in order to assimilate data it must only be ensured that the model can be used to calculate the "modeled" observations for any given model state. The modeled observations can then be compared to real data \( d \) to estimate the error made by the model. This error is quantified by some metric (introduced below) and subsequently minimized. The model state corresponding to minimum error is considered optimal and will be denoted \( x_{\text{opt}} \).

The motivation for improving our knowledge about the model state \( x_{\text{opt}} \) by combining observations and model can be twofold:

1. We want to constrain the observations \( d \), for example by interpolation of observational data with temporal gaps or limited spatial coverage. The modeled observations corresponding to propagating the optimized model state through the model, i.e. \( d_{\text{opt}} = Q(x_{\text{opt}}) \), then serve as improved observations.

2. We want to constrain the inputs to the model, or to reduce their uncertainties, by utilizing the measurements. That is, we are interested in the optimized model state itself.

The second application is the one for the estimation of \( \text{CO}_2 \) surface fluxes, because the fluxes are contained in the input vector \( x \). Correcting measurements of atmospheric \( \text{CO}_2 \) is not the primary goal in these studies. However, the comparison of the optimized observations \( d_{\text{opt}} \) with the actual measurements \( d \) can be used to assess the success of the optimization and to help identify biases (see for example figure 4.3 for such a check).

The most common way to quantify the error made by the model, or the model-data mismatch, is by measuring the distance between the observation \( d \) and the model output \( Q(x) \), i.e.

\[ J(x) = \| Q(x) - d \|^2. \]  

(2.2)

Here \( \| \cdot \| \) denotes a norm to be specified and \( J \) is the cost function, which we seek to minimize with respect to that norm. If prior information on the desired variables in the input
vector are available, they can be included in the cost function by similar means,

\[ J(x) = \alpha \cdot \| Q(x) - d \|^2 + \beta \cdot \| x - x_0 \|^2, \]  

(2.3)

where the zero index \( (x_0) \) denotes the prior values. Minimizing this cost function will result in a compromise between the information given by the observations and the information given by the prior values. The constants \( \alpha, \beta \) are the weight given to both terms. They should be chosen to represent the confidence in the observations and the prior information, respectively.

The determination of the distance norm and, hence, the definition of the cost function is a key task within the optimization procedure. It can be difficult to define cost function terms, particularly if the processes underlying the problem are not well understood. For example, it appears straightforward to weight each measurement in the observation vector by its inverse measurement error, but then the prior uncertainty must be chosen appropriately to balance both terms in the cost function. One diagnostic that plays a role in this regard is the \( \chi^2 \) statistics, as mentioned below. It is often desirable to take error correlations into account as well. This leads to the most commonly used form of the cost function taking into account the error covariance matrix:

\[ 2J(x) = (Q(x) - d)^T C_d^{-1} (Q(x) - d) + (x - x_0)^T C_0^{-1} (x - x_0). \]  

(2.4)

The factor 2 has no influence on the minimization results; I added it here already because it arises when the PDF (Probability Density Function) of both the observations and the prior constraint are Gaussian, as shown below. Both terms in the cost function are weighted by the inverse covariance matrices of the observational data \( (C_d) \) and the prior \( (C_0) \). A covariance matrix carries the variances (i.e. squared uncertainties) on its main diagonal and the correlations in its off-diagonal elements.

The Bayesian perspective

The same general form of the cost function is obtained from a Bayesian point of view, if all underlying probability densities are assumed Gaussian. I outline here the Bayesian idea as an additional motivation for the specific choice of distance metric that was implicitly taken above to proceed from equation (2.3) to (2.4).

The Bayesian concept is the ideal probabilistic framework for combining information from different sources. It provides a complete and general perspective on the problem of data assimilation and links many approaches to the problem in a simple and convincing way. The first step in Bayesian inference is to define the so-called full probability model, which represents the joint PDF of all parameters in both data and model space (e.g. observations \( d \), model parameters such as surface fluxes \( x \)). Then the conditional probability distribution
of the parameters of interest given the observed data is sought. According to Bayes’s rule, the joint PDF for \(x\) and \(d\), i.e. \(P(x, d)\), can be written in two ways:

\[
P(x, d) = P(x \mid d) \cdot P(d) = P(d \mid x) \cdot P(x) .
\]

\(^{(2.5)}\)

\(P(x)\) is the prior knowledge on \(x\), quantifying the prior understanding of the parameters of interest. The conditional probability of the data given the input parameters, \(P(d \mid x)\), is determined from \(x\) and the observation operator (i.e. the model). It is the probability distribution of the modeled data, after the model state \(x\) was propagated through the model. The key term in the above equation is \(P(x \mid d)\), as it represents the \textit{a posteriori} (henceforth, posterior) conditional probability distribution of a certain model state given the observations. Maximizing this probability with respect to the model state will provide the most probable model state as constrained by observations and, therefore, an optimal estimation of the parameters of interest, \(x_{\text{opt}}\). No specific conditions have to be met for any of the parameter’s PDF, except that the PDFs must be known and all of the conditional probabilities can be generated (see Tarantola (2005a) for a more detailed discussion on the statistical requirements).

Often the data distribution \(P(d)\) is statistically independent of the model state, so it can be viewed as a normalization constant only, thus simplifying equation (2.5) to

\[
P(x \mid d) \propto P(d \mid x) \cdot P(x) .
\]

\(^{(2.6)}\)

This is valid for the atmospheric and oceanic data distributions in all my inversions. If one considers that the PDFs have a Gaussian shape, one can write:

\[
P(x) = \exp \left( -\frac{1}{2} (x - x_0)^T C_0^{-1} (x - x_0) \right)
\]

\(^{(2.7)}\)

\[
P(d \mid x) = \exp \left( -\frac{1}{2} (Q(x) - d)^T C_d^{-1} (Q(x) - d) \right).
\]

\(^{(2.8)}\)

The optimal model state \(x_{\text{opt}}\) is derived by maximizing the posterior probability \(P(x \mid d)\). Inserting (2.7) and (2.8) into (2.6) and taking the logarithm leads to

\[
-\log P(x \mid d) \propto \frac{1}{2} \left( (Q(x) - d)^T C_d^{-1} (Q(x) - d) + (x - x_0)^T C_0^{-1} (x - x_0) \right) = J(x).
\]

\(^{(2.9)}\)

Because probabilities are represented by numbers between zero and one, i.e. \(P \in [0, 1]\), maximizing \(P(x \mid d)\) is equivalent to minimizing \(-\log P(x \mid d)\) and, hence, the cost function \(J(x)\). The cost function derived here is identical to the one introduced above, though following directly from the Bayesian concept in conjunction with Gaussian PDFs. Figure 2.1 shows a schematic of the minimization of a quadratic cost function. A cost function of the form (2.4)
would be quadratic if the model $Q$ is linear, as is the case in my atmospheric and oceanic inversions below.

![Figure 2.1 Minimization of a quadratic cost function $J(x)$, defined over a two-dimensional parameter space. The transition from the prior estimate $x_0$ to the optimized posterior estimate $x_{opt}$ is indicated. The projections of the cost function on the parameter plane are elliptic iso-cost lines with semi-axis lengths determined by the gradient of the cost function and related to the uncertainty and error covariance of the optimal parameter estimate.]

### 2.2 Atmosphere-only CO$_2$ inversion

The minimization of the cost function in (2.4) can be difficult in general, depending on the dimensionality of the problem and the complexity of the model $Q$. For my atmospheric inversion I take advantage of the linearity of atmospheric transport with regard to passive tracers such as CO$_2$, which allows for an explicit minimization scheme via matrix inversion. In addition, no prior CO$_2$ fluxes will be used, which further simplifies the problem and allows for fast solution even when implemented in a scripting language such as Matlab in my case. The statistical background I provide in this section does not take into account prior information, because the atmosphere inversion does not use explicit priors. The equations are, however, extended in section 2.5 in the framework of joining the atmosphere with the ocean inversion, where the joint inversion is a Bayesian synthesis method. The atmosphere-only inversion represents "mode 1" of my joint inversion.

#### 2.2.1 Linear synthesis inversion

The atmosphere inversion is a so-called synthesis method (Enting, 2002). In a first step the Earth’s surface is split into a set of regions defined based on latitude (oceanic regions) and vegetation type (land regions). The regional setup I used is identical to the one defined by Transcom (Gurney et al., 2002) and includes 11 oceanic and 11 terrestrial regions (figure
The CO$_2$ fluxes from these regions into the atmosphere are contained as parameters in the model state vector $x$. The next step is to release a dye (passive) tracer flux with prescribed strength and prescribed within-region flux pattern from each region. These regionally distributed fluxes represent basis functions (or Green’s functions), $v_r$, from which a global flux pattern $v$ is constructed by linear combination,

$$v = \sum_r x_r v_r, \quad (2.10)$$

where the index $r$ indicates the regions and the $x_r$ are the regional flux scaling factors, i.e. the elements of $x$. According to the data assimilation concept described above, the basic idea is to propagate this global flux pattern through an atmospheric transport model $Q$ and compare the resulting three-dimensional field of dye concentrations to an observed atmospheric concentration field $d$:

$$d = Q(v) = Q \left( \sum_r x_r v_r \right). \quad (2.11)$$

The task of the inversion is to solve for the scaling factors $x_r$. Because CO$_2$ is a passive tracer in the atmosphere, the model $Q$ is linear, that is, a doubling in the surface CO$_2$ release will result in doubled CO$_2$ concentrations at every point in the atmosphere, i.e. $Q(2v) = 2Q(v)$. Given the linearity of transport, $(2.11)$ simplifies to

$$d = \sum_r x_r Q(v_r). \quad (2.12)$$

Equation $(2.12)$ allows to repeatedly run the transport model on each basis function without the need to combine them first to a global flux pattern. The obtained regional response functions can then be scaled in postprocessing to "synthesize" the modeled atmospheric CO$_2$ concentration field.

The observed atmospheric CO$_2$ concentrations (the elements of the column data vector $d$) are not available as three-dimensional field, but as measurements at predefined times and places. Hence, the model output (right-hand side of $(2.12)$) is sampled at the coordinates of the observations. Owing to the linearity of transport we can sample the regional response functions $Q(v_r)$ individually and write the result in a column vector. The operation of scaling the response functions then translates into a matrix operation on the scaling factors $x_r$,

$$d_a = M_a x, \quad (2.13)$$

where $M_a$ is a matrix representing the “discretized” atmospheric transport and $d_a$ carries the atmospheric CO$_2$ measurements at the specified coordinates. In what follows I will speak
2.2. Atmosphere-only CO$_2$ inversion

Figure 2.2  Shown are in blue and green colors the 11 regions over ocean and land, respectively. Red points show the observation sites for atmospheric CO$_2$. White points correspond to the locations of ocean interior DIC profile measurements used in the ocean inversion underlying mode 2 (and all further modes) of our joint inversion. For the 1992-1996 inversion, monthly mean atmospheric CO$_2$ measurements from 85 sites and more than 60,000 ocean DIC observations are used. The decadal mean inversions use various atmospheric CO$_2$ networks, see figure 4.2 in chapter 4 for details.

of x as the regional fluxes, although they are strictly speaking only the scaling factors of the underlying flux patterns.

Equation (2.13) can now be solved using the concepts introduced in section 2.1, that is, by minimizing the mismatch between modeled concentrations $M_a x$ and observed concentrations $d$. The mismatch is expressed as a cost function according to (2.4),

$$J(x) = \frac{1}{2} (M_a x - d_a)^T C_{d_a}^{-1} (M_a x - d_a),$$

(2.14)

with the observational data covariance matrix $C_{d_a}$, which inversely weights the atmospheric observations. The name "data covariance" or even "data uncertainty" is widely used (henceforth, I will call it data covariance, too), though it is misleading, because the term must also account for the inability of the transport model (Tarantola, 2005a, chap.3), e.g. imperfect transport and coarse spatial and temporal resolution. In addition, the predefined within-region flux pattern or "footprint" for the basis functions may not represent the true distribution and, hence, may add to the model’s inability (Kaminski et al., 2001). In the inversion, it would be inappropriate to attempt to fit the observations better than the sum of all these errors, yet the quantification of this overall model-data mismatch in an objective manner remains difficult. Within the Transcom series of inversions the error due to imperfect model transport was assessed by repeating the inversions with different models and using the spread of results...
as an estimate for model transport uncertainty. This additional "between-model" uncertainty was added to the posterior flux uncertainty (more information on the error assessment follows in the following sections). I use the same approach by using the same 12 atmospheric transport models that participated in the Transcom study. However, using a suite of transport models does not rule out that all models may be biased in a structurally similar way.

The minimization of (2.14) yields

\[
\begin{align*}
    x_{\text{opt}} &= \left( M_a^T C_{d_a}^{-1} M_a \right)^{-1} M_a^T C_{d_a}^{-1} d_a, \\
    C_{\text{opt}} &= \left( M_a^T C_{d_a}^{-1} M_a \right)^{-1},
\end{align*}
\]

(2.15)

(2.16)

where \( x_{\text{opt}} \) and \( C_{\text{opt}} \) are the optimized flux estimate along with covariance matrix, thus representing the posterior PDF. \( C_{\text{opt}} \) is often called the internal or "within-model" covariance to distinguish it from the "between-model" covariance based on model spread. \( M_a^{-1} \) is the generalized inverse of \( M_a \); it is useful to illustrate the relationship of (2.15) with the solution of a system of linear equations: if in (2.13) the dimensions of the data and parameter vectors were the same and all observations and transport matrix elements were perfectly accurate (zero uncertainty), the equation would define such a system. It could be solved by multiplying the inverted (square) transport matrix with the data vector. Hence, in a least-squares sense, \( M_a^{-1} \) takes over the role of \( M_a^{-1} \) and justifies the common name "inversion" for the method. I compute the posterior PDF by direct calculation of \( x_{\text{opt}} \) and \( C_{\text{opt}} \) according to the above equations; the computation time in Matlab does not exceed a few minutes due to the reasonable size (generally less than 1 million elements) of the involved matrices.

The minimum value of the cost function, \( J_{\text{min}} = J(x_{\text{opt}}) \), reflects the degree to which the inversion calculation matches the data. It is usually expressed using the reduced \( \chi^2 \),

\[
\chi^2 = \frac{J_{\text{min}}}{N_{\text{obs}}} = \frac{1}{N_{\text{obs}}} \sum_{n=1}^{N_{\text{obs}}} \frac{(M_a x_{\text{opt}} - d_a)_n^2}{\sigma_n},
\]

(2.17)

where \( N_{\text{obs}} \) is the number of observations, that is, the length of the data vector \( d_a \). \( \sigma_n \) is the uncertainty (one standard deviation) of observation \( n \), derived as the square root of the \( n \)th element on the main diagonal of \( C_{d_a} \). Statistical consistency requires \( \chi^2 \) not to be significantly different from unity, otherwise the posterior uncertainty is inconsistent with the quality of the fit to the data. A value \( \chi^2 > 1 \), for example, suggests that too much confidence has been put in the ability of the inversion to match the data, so that the posterior model-data mismatch is greater than the prior data uncertainty. In such a case one would increase the prior data uncertainty accordingly to get the \( \chi^2 \) value closer to one. This would in turn increase the posterior flux uncertainty \( C_{\text{opt}} \), as can be seen from (2.16). In all my inversions I apply an iterative procedure to achieve \( \chi^2 \simeq 1 \) when averaged over all transport models.
2.2. Atmosphere-only CO$_2$ inversion

My approach for the atmospheric CO$_2$ inversion has a temporal resolution of months, though not interannual but averaged over inversion periods of several (5 to 10) years. Following Gurney et al. (2004), I call it a monthly cyclostationary inversion. Monthly CO$_2$ observations are combined with 11 or 12 of the Transcom atmospheric transport models to estimate the cyclostationary air-sea and air-land fluxes of carbon from each of the 22 Transcom regions. Given 22 regions and 12 months, the total number of model parameters is $12 \times 22 = 264$. In the Transcom models, the strength of the regional CO$_2$ release (basis functions) was prescribed to $1$ Pg C yr$^{-1}$. The within-region flux patterns were prescribed with model-based monthly biospheric CO$_2$ flux (for the land regions) and surface ocean pCO$_2$-based air-sea flux estimates. The patterns are shown in figure 2.3 averaged over the boreal summer and winter months. My inversion setup is similar to that of Gurney et al. (2004), but with some important deviations. A major difference is that I do not use prior flux estimates to regularize the posterior flux estimates. Other deviations include the construction of the observational network, the assignment of observational (prior) errors and the global fossil fuel emission estimates. In the following sections I describe the setup briefly and focus on the differences. For more details on the Transcom inversion setup I refer to Gurney et al. (2002), Transcom 3 level 1 (T3L1), and Gurney et al. (2004), Transcom 3 level 2 (T3L2), as well as the Transcom experimental protocol.

2.2.2 Network choice and data handling

Atmospheric CO$_2$ data to be inverted stem from the recent 2009 version of the NOAA Globalview CO$_2$ dataset (Globalview-co2, 2009). The dataset contains 283 stations with measurements made between 1979 and 2009 (last fully covered year is 2008). Data do not represent direct measurements, but are the result of a smoothing and inter-/extrapolating procedure. At the base of this procedure is a latitudinal boundary layer reference CO$_2$ dataset in conjunction with site-specific baseline selection (see Masarie and Tans (1995) for more details). As a result, the data from all sites are available with weekly resolution and cover the whole 1979-2008 period. While this makes the handling of these data easy, especially in an inverse modeling context, it must be kept in mind that many stations collected actual measurements only for short time periods, so that a large fraction of the provided data solely represent reference values. To take this into account we considered only those stations, which fulfill the two selection criteria:

1. The total number of real measurements throughout the inversion period must be at least 50% (for the decadal inversions between 1980 and 2008) or 70% (for the 1992-1996 inversion) of all data.

2. For each month the fraction of real measurements must be at least 40%, when aver-
aged over the inversion period.

The first criterion follows the T3L2 scheme; it counts all real measurements during the inversion period and makes sure that their number is at least as high as that of processed data. The reason for relaxing the threshold from 70% to 50% in the decadal inversions is described in section 4.3.1. The second criterion ensures that there is no severe seasonal bias in the fraction of real data, i.e. that certain months are not over- or under-represented compared to others. For example, this could happen for stations, which sampled regularly during summer, but only rarely during winter. As in the Transcom inversions, we explicitly exclude observations from the WITN surface and Darwin stations owing to concerns about data quality and local representativeness (Gurney et al., 2003; Law et al., 2003).

From the weekly Globalview data we first derive monthly mean CO$_2$ concentrations for each month in the inversion period. These are then detrended using a fit that includes a linear trend as well as harmonics. Finally, a cyclostationary seasonal cycle is computed by averaging the data over the inversion period by month. These decadal mean, monthly data are used to constrain the inversion.

The uncertainty assigned to each data point is mainly based on the monthly residual standard deviation (RSD$_{\text{mon}}$) of the direct measurements around the smoothed and extended Globalview data series as well as the monthly and total ratios of real measurements to processed data ($P_m$ and $P_t$, respectively). It is calculated as

$$
\sigma_{\text{mon}} = \max \left[ \sigma_{\text{mon}}^{\text{min}} ; \frac{\text{RSD}_{\text{mon}}}{\sqrt{A \cdot P_t \cdot P_m}} \right] \cdot \sqrt{N_{\text{coloc}}},
$$

(2.18)

with $\sigma_{\text{mon}}$ representing the square root of each element on the main diagonal of the inversion's prior data covariance matrix $C_d$ (as introduced in the previous section). $\sigma_{\text{mon}}^{\text{min}}$ is a minimum monthly uncertainty, which is inferred from a corresponding minimum condition for the annualized uncertainty, following the T3L2 procedure. Briefly, an annual uncertainty is computed on the basis of the monthly values and the station's auto-correlation timescale. If the annual uncertainty is below a minimum value of 0.25 ppmv, the uncertainty for each month is set to this value multiplied by the auto-correlation timescale. As a result, $\sigma_{\text{mon}}^{\text{min}}$ is a site-specific constant. However, as the auto-correlation timescale of most sites is around 3 months, $\sigma_{\text{mon}}^{\text{min}}$ takes values around 0.43 ppmv and does not deviate much from site to site. Note that monthly uncertainties smaller than this value are still possible if the corresponding annualized uncertainty condition is fulfilled (figure 4.3). $N_{\text{coloc}}$ indicates the number of co-located sites, i.e. sites which are located within the same grid cell of the average model grid. To avoid double-counting, data from those sites are de-weighted. The constant $A$ is chosen iteratively to achieve a mean reduced $\chi^2$ close to 1. While the consideration of $P_m$ generally increases the uncertainty estimate, this is on average compensated by the adjustment of the constant $A$ to meet the $\chi^2$ condition. As a result, the overall level of uncertainty is very
2.2. Atmosphere-only CO$_2$ inversion

similar to T3L2, but stations with low real data coverage for some months are de-weighted compared to stations with continuously high monthly coverage.

Once an inversion period is chosen, I construct an individual network according to the described scheme. For the decadal inversions in chapter 4 I constructed additionally a common network covering the decades between 1980 and 2008. For details about how the common network is constructed, see section 4.3.1.

2.2.3 Background flux fields

Prior to inversion, we effectively remove from the data the atmospheric CO$_2$ response to emissions from fossil fuel burning, cement manufacture and gas flaring. This response is computed by propagating CO$_2$ emission distributions for 1990 (Brenkert et al., 1997) and 1995 (Brenkert et al., 1998) through the atmospheric transport models, followed by sampling the concentration response at data locations and months. The 1995 distribution is shown in figure 1.5. After scaling the contributions from both reference years to approximate the mean emissions over the inversion period, they are prescribed in the inversion with very low uncertainty. While for the T3L2 inversions, the scaled global integrals from the 1990 and 1995 emission maps were used to obtain global total emissions for the 1992-1996 period, we use annual emission rates from the dataset of Marland et al. (2008). This allows us to keep global emissions consistent with annual estimates when extending our inversion approach to other periods. For the 1992-1996 period, fossil CO$_2$ emissions sum up to 6.3 Pg C yr$^{-1}$ globally, which is slightly higher than the value of 6.1 Pg C yr$^{-1}$ used by T3L2. For other periods the distribution of fossil fuel emissions is inferred from the 1990 and 1995 reference fields according to the following scheme: for years before 1990, the 1990 reference field is scaled to match the global rate of Marland et al. (2008); the 1995 reference field is not used. For years between 1990 and 1995, both reference fields are scaled by the relative distance from the specified year, with the global rate again adjusted to Marland et al. (2008). For years after 1995, the 1995 reference field is scaled to match the global rate of Marland et al. (2008); the 1990 reference field is not used.

In addition to this fossil emissions background field, two more background fields are subtracted from the data prior to inversion: an annually balanced, seasonal terrestrial biosphere exchange and air-sea gas exchange. The terrestrial biospheric exchange pattern is the same as used for prescribing the flux patterns within the regional basis functions (figure 2.3). The flux pattern is purely seasonal (annual mean flux is zero for every pixel). The presubtracted air-sea flux is based on the same pCO$_2$ data as used for the basis functions. The annually integrated presubtracted air-sea flux is -2.19 Pg C yr$^{-1}$ (uptake by the ocean). These background fields were propagated through the transport models, just like the fossil fuel background field (CO$_2^{FF}$), and the modeled response was sampled at the coordinates of
Figure 2.3  Global distribution of NEP (Net Ecosystem Production) for land regions and net carbon flux across the air-sea interface (positive fluxes are directed into the atmosphere). The upper panel shows the distribution averaged over the boreal winter (here October to March), while boreal summer (here April to September) is shown in the lower panel. Air-sea fluxes are on the same scale as terrestrial biospheric fluxes to point out the terrestrial dominance in the seasonal signal. These distributions are used in the synthesis inversion described in the text to infer regional flux patterns. The seasonal biospheric exchange is based on the Carnegie Ames Stanford Approach (CASA) model (Randerson et al., 1997) and has an annual flux of zero. The oceanic exchange is based on pCO$_2$ measurements of Takahashi et al. (1997).

the observations and subsequently removed from the observational data. As a result, the inversion estimates carbon fluxes to fit CO$_2$ concentration anomalies (CO$_{2\text{anom}}$) rather than absolute observed CO$_2$ concentrations (CO$_{2\text{obs}}$). With the biospheric background response
(CO$_2^{NEP}$) and the air-sea background response (CO$_2^{AS}$), the anomalous concentration writes

$$CO_2^{anom} = CO_2^{obs} - CO_2^{FF} - CO_2^{NEP} - CO_2^{AS}. \quad (2.19)$$

It should be noted that the background flux fields do not represent prior information (they operate in data space). The flux estimates resulting from the inversion of the anomalous concentration field implicitly include corrections to these assumed background flux fields.

During the inversion the mismatch between the simulated data and observed CO$_2$ concentration anomalies is minimized by least squares through optimization of the flux contributions from each region and month. The covariance matrix provided along with these flux estimates, i.e. flux uncertainty and error correlation, represents the "within-model" covariance. A "between-model" covariance matrix is also computed from the spread of results among all transport models, represented as one standard deviation from the model mean. These two contributions are summed to obtain the "total" covariance matrix. Reported flux uncertainties will always represent the square root of the diagonal elements of this total covariance matrix, unless otherwise noted.

### 2.2.4 Global CO$_2$ growth constraint

In addition to the observational data, the global carbon uptake by the oceans and the terrestrial biosphere constrains the inversion. It is derived as the difference between global fossil fuel emissions (including emissions from cement production and gas flaring) and observed atmospheric CO$_2$ growth rate. The latter is estimated as the average of the linear trends of CO$_2$ concentration at the Mauna Loa and South Pole stations. Trends are obtained by fitting a linear model with three harmonics (with periods of 12, 6 and 4 months, respectively) to the Globalview data of the two stations. I consider half the difference between the two estimates as growth rate uncertainty. For the T3L2 inversions, the mean trend of all 75 stations was used, with their standard deviation as uncertainty estimate. However, the global distribution of the 75 sites is very inhomogeneous, with most stations located in the Northern Hemisphere land regions (in particular in Europe and North America), a few in the Southern Hemisphere and almost none in the Tropics. The growth rate therefore may be biased towards the Northern Hemisphere and the uncertainty may be underestimated due to the high density of stations in regions with similar CO$_2$ signal and thus similar trend. To avoid such a potential bias towards the better-sampled Northern Hemisphere, I decided to consider only one station for each Hemisphere. For example, for the 1992-1996 period I derive an atmospheric growth rate of 3.18 (±0.16) Pg C yr$^{-1}$, as opposed to the Transcom estimate of 3.27 (±0.07) Pg C yr$^{-1}$. As a result, global uptake is increased by about 10% in my inversion, i.e. 3.12 (±0.16) Pg C yr$^{-1}$ compared to 2.83 (±0.07) Pg C yr$^{-1}$ for the Transcom setup.
Chapter 2. Inverse method

2.3 Ocean interior DIC inversion

My ocean inversion is methodologically identical to the one first developed by Gloor et al. (2003), then expanded and refined by Mikaloff Fletcher et al. (2006, 2007), and revisited by Gruber et al. (2009). I illustrate the method here, but refer to those publications for more details.

Like the atmosphere inversion, the ocean inversion uses the Green’s function or synthesis approach. After partitioning the global ocean surface into 30 regions, for each region a footprint (the response function) is simulated using an ocean circulation model. Each regional footprint represents the dye tracer concentration field throughout the global ocean resulting from injection of a unit flux of this tracer into one region. The footprints are then scaled in such a way that their weighted sum most closely matches a set of available observations. The scaling factors are obtained by least-squares minimization of the mismatch between observations and modeled responses, with the latter being sampled at the locations of the observations. In figure 2.4 the regional footprints and the sampling procedure are illustrated. Clearly distinct footprints are seen, particularly between the three major ocean basins. Using the concepts introduced in section 2.1, the scaling factors $x$ representing regional air-sea fluxes are estimated according to equation (2.14) by minimizing the cost function

$$J(x) = \frac{1}{2} (M_o x - d_o)^T C_{d_o}^{-1} (M_o x - d_o) .$$

The transport matrix $M_o$ represents the “discretized” ocean circulation model, that is, it operates on the flux vector $x$ and samples the modeled response at the locations of the observations $d_o$. $C_{d_o}$ is the diagonal ocean carbon data covariance matrix carrying the data uncertainties on its main diagonal. The posterior PDF, consisting of the optimized regional flux vector $x_{opt}$ and covariance matrix $C_{opt}$, is derived analogous to equations (2.15) and (2.16) by replacing the atmosphere subscript $a$ with the ocean subscript $o$.

As detailed below, two different tracers are separately inverted, one representing the concentration of natural “preindustrial” carbon in the ocean and the second the portion of carbon that has arisen from the anthropogenic perturbation of atmospheric CO$_2$ since 1765, defined here as the onset of industrialization. The choice of size and location of each region was guided by the conception of minimizing flux heterogeneity within the region (Gloor et al., 2003), where estimates of regional flux heterogeneity are based on $\Delta p$CO$_2$ patterns of Takahashi et al. (1997).

Ocean interior observations to be inverted are available to a large extent in the GLODAP version 1.0 database (Key et al., 2004), augmented with some data collected during historical cruises in the Atlantic (Mikaloff Fletcher et al., 2006). The database contains observations of many oceanographic parameters, sampled at stations distributed quite homogeneously.
2.3. Ocean interior DIC inversion

Figure 2.4 Global zonal mean response functions or footprints for the 11 Transcom oceanic regions. The colored pattern is the quasi-steady state dye tracer concentration in response to the injection of a continuous flux of 1 Pg C yr$^{-1}$ into each region. Units are arbitrary, only the spatial tracer distribution matters for the inversion. Black dots represent measurement locations (only schematically), where the response functions are sampled; these "discretized" responses form the ocean transport matrix $M_o$ in the cost function.

throughout the global ocean (figure 2.2). At each station the whole water column was generally sampled, resulting in a total number of observations exceeding 68,000. This high data density, especially when compared to the sparse observational network for atmospheric CO$_2$ (section 2.2), is responsible for the strong over-determined character of the ocean inversion, resulting in small internal uncertainties of the inverse flux estimates. It also allows to use a relatively big number (when compared to the atmospheric inversion) of 30 regions. A more extensive set of regions reduces aggregation errors (Kaminski et al., 2001) associated with inversions for large-scale regions. While most of the observations were taken in the 1990s, they do not solely represent this decade. Sampling the whole water column leads to an effective temporal smoothing of data, as water volumes below the mixed layer consist of water masses with different origins and thus different age. Observations, therefore, represent carbon concentrations averaged over the time scales of ocean circulation. Air-sea CO$_2$ fluxes estimated by an inverse method consequently constitute long-term mean fluxes and do not provide interannual information, nor do they resolve the seasonal cycle.
2.3.1 Preindustrial and anthropogenic carbon tracers

The ocean inversion does not estimate regional fluxes of CO$_2$ based on direct observations of dissolved inorganic carbon (DIC), but separately inverts data-based inventories of two quasi-conservative tracers: anthropogenic ($\Delta C_{ant}$) and natural "preindustrial" carbon ($\Delta C_{gasex}$). The anthropogenic CO$_2$ is separated from the measured DIC based on the $\Delta C^*$ method of Gruber et al. (1996). For detailed information about the computation of $\Delta C_{ant}$ the reader is referred to, e.g., Gruber (1998). In summary, the anthropogenic component of DIC within a water parcel is estimated in two steps from the parcel’s history since it was last in contact with the atmosphere. In the first step the biological imprint on DIC, $\Delta C_{bio}$, is removed. This biological term corrects for the addition of DIC due to remineralization processes and dissolution of CaCO$_3$. It can be estimated from data with the assumption of constant stoichiometric ratios P:N:C:O$_2$ between phosphate, nitrogen, organic carbon and oxygen (Anderson and Sarmiento, 1994). As a second step, the background preindustrial DIC concentration is subtracted. This term is formally split into two components, the preindustrial equilibrium concentration, DIC$_{eq}$, and the preindustrial disequilibrium concentration, $\Delta C_{diseq}$. The anthropogenic component of the measured DIC can then be computed as

$$\Delta C_{ant} = DIC - \Delta C_{bio} - DIC_{eq} - \Delta C_{diseq} = \Delta C^* - \Delta C_{diseq},$$

where $\Delta C^*$ is defined as the sum of the first three terms. DIC$_{eq}$ represents the concentration of DIC in a water parcel equilibrated with the preindustrial atmosphere and can be calculated based on hydrographic observations. The estimation of $\Delta C_{diseq}$, i.e. the deviation from the equilibrium concentration, additionally requires knowledge about the water age, defined as the time between a water parcel lost contact to the atmosphere and its sampling. Water age can be determined by, e.g., CFC measurements (chapter 1). Resulting $\Delta C_{ant}$ concentration fields for several ocean basins are shown in figure 2.5; see also Mikaloff Fletcher et al. (2006). Sabine et al. (2004) give a comprehensive summary about anthropogenic carbon in the global ocean.

The component of DIC reflecting preindustrial gas exchange is represented by the tracer

$$\Delta C_{gasex} = DIC - \Delta C_{bio} - \Delta C_{ant} - DIC^0,$$

where DIC$^0$ is an arbitrary constant chosen such that the average $\Delta C_{gasex}$ within the global surface ocean is close to zero. This offset is not of relevance for the inversion, as the inversion interprets spatial gradients in concentration only. For computational details we refer to Gruber and Sarmiento (2002). By definition, changes in this tracer can only be caused by preindustrial gas exchange combined with oceanic transport and mixing. Representative illustrations of $\Delta C_{gasex}$ are shown in figure 2.5; see also Mikaloff Fletcher et al. (2007).
Both the anthropogenic and the preindustrial carbon tracers are quasi-conservative, i.e. there are no sinks and sources of $\Delta C_{ant}$ and $\Delta C_{gasex}$ in the ocean interior. Given that we know how the ocean circulation operates, we can thus reconstruct the strength and distribution of CO$_2$ exchange fluxes across the air-sea interface that have led to the "observed" concentration fields of the two tracers.

![Concentration profiles of ocean carbon tracers along North-South transects in the Pacific (left panel) and Atlantic (right panel) ocean basins.](image)

The rows represent the carbon concentration terms as defined in the text and in equation (2.23). From top to bottom they are observed DIC, the biological correction term ($\Delta C_{bio}$), the anthropogenic carbon concentration ($\Delta C_{ant}$), and the preindustrial gas exchange tracer ($\Delta C_{gasex}$). Selected isopycnal surfaces (in kg m$^{-3}$) are contoured in the bottom row. Black dots in the top row show the locations of the DIC observations from which these fields were interpolated; their latitudinal positions are very similar to the profile locations depicted in figure 2.2. Thick vertical lines represent the regional boundaries defined within the Transcom project, cf. blue regions in figure 2.2. Thin vertical lines show the boundaries from 30 regions as defined in the ocean inversion, before they were aggregated to the 22 Transcom regions. Figure from Jacobson et al. (2007b).

**Figure 2.5** Concentration profiles of ocean carbon tracers along North-South transects in the Pacific (left panel) and Atlantic (right panel) ocean basins. The rows represent the carbon concentration terms as defined in the text and in equation (2.23). From top to bottom they are observed DIC, the biological correction term ($\Delta C_{bio}$), the anthropogenic carbon concentration ($\Delta C_{ant}$), and the preindustrial gas exchange tracer ($\Delta C_{gasex}$). Selected isopycnal surfaces (in kg m$^{-3}$) are contoured in the bottom row. Black dots in the top row show the locations of the DIC observations from which these fields were interpolated; their latitudinal positions are very similar to the profile locations depicted in figure 2.2. Thick vertical lines represent the regional boundaries defined within the Transcom project, cf. blue regions in figure 2.2. Thin vertical lines show the boundaries from 30 regions as defined in the ocean inversion, before they were aggregated to the 22 Transcom regions. Figure from Jacobson et al. (2007b).

Both the anthropogenic and the preindustrial carbon components do not represent direct measurements, but are data-derived quantities. Their computation can be expressed by
viewing at their functional dependencies,

\[
\Delta C_{\text{ant}} = f(T, S, DIC, Alk, AOU; r_{C:O}, r_{N:O}; \Delta C_{\text{diseq}}),
\]

(2.24)

\[
\Delta C_{\text{gasex}} = f(S, DIC, PO_4^{3-}, Alk; r_{C:P}, r_{N:P}; \Delta C_{\text{ant}}).
\]

(2.25)

To assign observational uncertainties, one not only has to propagate random errors in the measurements of the underlying hydrographic parameters, but also to account for potential biases related to the choice of stoichiometric ratios and the calculation of \(\Delta C_{\text{diseq}}\). \(\Delta C_{\text{gasex}}\) additionally inherits the uncertainty of \(\Delta C_{\text{ant}}\). For the anthropogenic component we adapt the uncertainty estimation scheme developed by Mikaloff Fletcher et al. (2006), based on earlier work of Gruber (1998). The preindustrial uncertainty scheme follows Mikaloff Fletcher et al. (2007), who applied a suite of sensitivity analyses to assess the impact of uncertain C:P and N:P stoichiometry. Typical values for the final uncertainty are 18 \(\mu\)mol kg\(^{-1}\) for \(\Delta C_{\text{ant}}\) and 29 \(\mu\)mol kg\(^{-1}\) for \(\Delta C_{\text{gasex}}\).

The simulation setup for the preindustrial and anthropogenic footprints differ from each other.

To simulate the preindustrial footprint for each region, a continuous unit flux of dye was injected into the surface ocean in this region. The spatial flux pattern within the region was prescribed based on the climatology of Takahashi et al. (1997). The aim was to then run the models until a quasi-steady state is reached, i.e. until the dye concentration changes at the same rate everywhere in the global ocean. Though this was not perfectly achievable, it turned out that a simulation time of about 3000 years is sufficient for that purpose. When normalized, this quasi-steady state dye distribution represents the regional footprint of preindustrial carbon. Conversely, the simulation of an anthropogenic footprint involves the application of a time-varying surface flux corresponding to the increase in atmospheric CO\(_2\) over the anthropogenic era, i.e. 1765-1999 in this study. The response in ocean surface flux was assumed to scale linearly with the atmospheric CO\(_2\) perturbation, following Mikaloff Fletcher et al. (2006). This assumption holds in case of an exponentially increasing atmospheric CO\(_2\) burden in conjunction with a constant oceanic buffer factor (Sarmiento et al., 1995). For the 1990s, when most of the DIC measurements were made, both conditions were approximately fulfilled. By design of these footprint simulations, inverse results for anthropogenic CO\(_2\) fluxes always reflect the regional flux contributions averaged over the whole anthropogenic era. In order to reference anthropogenic flux results to a specific year, they are scaled by that year’s atmospheric CO\(_2\) perturbation, cf. figure 4.5.

Both preindustrial and anthropogenic air-sea CO\(_2\) fluxes are initially estimated for 30 regions, but aggregated in postprocessing to 11 regions matching the Transcom regions used for the atmosphere inversion. To aggregate the fluxes I mapped the scaling factors back on the predefined pCO\(_2\)-based regional flux patterns to obtain a two-dimensional flux map. This flux map was then integrated over the Transcom regions to obtain the corresponding scaling factors. A small (<3%) global regridding error occurred due to the different grids of the pCO\(_2\)
dataset and the Transcom regions mask file. I distributed the deviation among the regions in proportion to their area, in order to maintain the global total flux estimate. To aggregate the covariance matrix I summed the Gaussian contributions to each Transcom region from those original regions that combine to the best approximation of this Transcom region. Because the original 30 regions are not a superset of the 11 Transcom regions, this approximation is not perfect. However, the latitudinal boundaries of both region sets are quite similar, so that the error is likely to be small, however difficult to quantify.

### 2.3.2 Net air-sea flux

A contemporary flux estimate is formed by summing preindustrial and anthropogenic flux components. While the contemporary flux represents the net CO$_2$ flux across the air-sea surface, the ocean interior dye distribution formed by the corresponding composition of simulated footprints does not represent contemporary DIC because the biological correction term as well as the background DIC$^0$ are still subtracted. This can readily be seen by adding $\Delta C_{\text{ant}}$ to equation (2.23). A detailed analysis of this contemporary air-sea CO$_2$ flux as well as implications for global oceanic transport patterns were compiled by Gruber et al. (2009).

The anthropogenic component is referenced to the midpoint of the time period of the atmosphere inversion before summing, so that the ocean inversion-based air-sea flux estimates represent the same period like the atmosphere inversion-based air-sea fluxes.

When combining an ocean with an atmosphere inversion, the riverine carbon loop needs to be handled consistently, i.e. the fraction of carbon that is taken up over land, then reaches a river and is transported into the ocean, where it is eventually released to the atmosphere. We adopt the perspective of Jacobson et al. (2007a) of an ideal atmosphere inversion that correctly represents this loop by assigning a land sink and corresponding ocean source of carbon. The ocean inversion, on the other hand, does not “see” the land, and misinterprets the increase in open ocean DIC, caused by riverine carbon input, as input across the sea surface. This leads to an overestimation of the preindustrial portion of the ocean sink that needs to be corrected for, before ocean inverse results are used in a joint inversion. This argument only holds for riverine carbon reaching the open ocean (cf. discussion in auxiliary material of Jacobson et al. (2007a)), as coastal regions are not resolved by the coarse resolution transport models, neither are they covered by observations used for the inversion. The riverine carbon input into the open ocean is resolved spatially according to Jacobson et al. (2007a) and sums up globally to 0.45 Pg C yr$^{-1}$. The correction is applied to posterior ocean inverse fluxes, leading to the correct attribution as land sink. The same value is used for all joint inversion periods.

My ocean inversion scheme is very similar to that of Jacobson et al. (2007a), with slightly
different error weighting scheme and an extended transport model set. While we use all 10 oceanic transport models that participated in the Ocean Inversion Project (Mikaloff Fletcher et al., 2006), they used a subset of five. The resulting differences in air-sea flux are very small, likely due to the fact that their model set actually consists of five different configurations of one transport model chosen to represent "fundamental axes" in transport uncertainty. Hence the addition of five transport models does not necessarily increase model diversity.

Due to the over-determined character of the ocean inversion, resulting flux uncertainties are very small, typically around 0.01 Pg C yr\(^{-1}\), less than 3% of the flux, for all transport models. Compared to this internal ("within-model") uncertainty, the spread among models is dominant in all regions. When combined as squared sum, the total uncertainty is typically 0.06 Pg C yr\(^{-1}\) or 22% of the flux, for the 11 regions. Due to mostly negative correlations the uncertainties for the aggregated 11 Transcom regions are smaller than the original uncertainties for 30 regions. This generally makes the "between-model" uncertainty dominant in all regions. The over-determined character also makes the use of a global sink constraint (as was used in the atmosphere inversion) obsolete. Results from the ocean inversion are strong constraints for annual air-sea CO\(_2\) fluxes averaged over a multi-decadal timescale. But they do not provide information on seasonal flux variability, which is why I augmented them with seasonal flux patterns from pCO\(_2\) measurements to produce a seasonally resolved oceanic constraint that can then be used in a joint inversion. How this composite constraint is constructed will be explained in the next few sections.

### 2.4 Surface ocean pCO\(_2\)-based flux information

Results from the ocean inversion represent the mean annual air-sea flux over, at least, several decades. The sluggish character of ocean circulation as well as the temporally smoothed nature of DIC measurements prevent the estimation of fluxes on shorter timescales. To estimate seasonal CO\(_2\) fluxes in the frame of a joint inversion, we extract seasonal flux patterns from the recent pCO\(_2\) climatology of Takahashi et al. (2009a) and map them on the annual inverse fluxes from the 11 regions.

In their database, Takahashi et al. (2009a) compiled more than 3.5 million observations of surface water pCO\(_2\), measured between 1968 and 2006. For the climatology (see figure 2.6), all measurements are referenced to the year 2000 by considering the mean pCO\(_2\) increase over the global surface ocean of 1.5 \(\mu\)atm yr\(^{-1}\). This value is based on fitting a linear trend to the deseasonalized pCO\(_2\) observations. For more information on this procedure as well as on their spatiotemporal interpolation scheme, we refer to Takahashi et al. (2009b). The database also contains atmospheric pCO\(_2\) observations for the year 2000, based on Globalview-co2 (2006). The difference of the two, \(\Delta pCO_2 = pCO_2(\text{sea water}) - pCO_2(\text{air})\), is the main driver
2.4. Surface ocean pCO$_2$-based flux information

Figure 2.6  Dataset of Takahashi et al. (2009b). The top picture show the annual mean, net air-sea flux climatology for a specific gas exchange model, based on more than 3 million pCO$_2$ measurements globally. Measurement locations (cruise tracks) are shown below. In the lower right the monthly observational density is shown, separated into observations from the Northern and Southern Hemispheres.

for air-sea CO$_2$ flux, with a positive value indicating a flux from the ocean into the atmosphere and vice versa. Computing a flux $F$ from $\Delta$pCO$_2$ involves the gas transfer velocity $k$ and the solubility $\alpha$,

$$ F = k \alpha \Delta pCO_2. $$

(2.26)
Chapter 2. Inverse method

The solubility depends on temperature and salinity and can be calculated precisely based on the parameterization of Weiss (1974). The gas transfer velocity is usually parameterized as a power function of windspeed, \( k \sim u_{10}^n \), where \( u_{10} \) is wind speed at 10 m elevation. The exponent \( n \) ranges from 1 (Smethie et al., 1985) to 3 (Wanninkhof and McGillis, 1999), with most studies supporting \( n = 2 \) (Wanninkhof, 1992; Nightingale et al., 2000; Ho et al., 2006; Sweeney et al., 2007).

The various parameterizations constitute a significant source of uncertainty for CO\(_2\) flux estimations derived from \( \Delta p_{CO_2} \) measurements. This can be seen in table 2.1, where we compare the global oceanic CO\(_2\) sink strength resulting from the various parameterizations. To account for this, we computed fluxes for eight different gas transfer velocity parameterizations and took the spread among them as a component of flux uncertainty. For each grid cell and month of the Takahashi et al. climatology we adopt the mean flux value of the eight parameterizations as best estimate. These monthly fluxes are then aggregated to the 11 regions of the inversion and mapped on the annual inverse estimates, i.e. for each region the fluxes for each month are offset by a constant flux so that their annual mean equals the inverse flux. The various error sources lead to less well-constrained absolute fluxes compared to the inverse estimates. However, we believe the seasonal \( \Delta p_{CO_2} \) based flux pattern to be robust. This is because the various gas transfer models, despite representing an important source of uncertainty, all imply a monotonic relationship between wind speed and CO\(_2\) flux, and therefore correctly translate gradients in wind speed into flux gradients.

In their error analysis, Takahashi et al. (2009b) focus on the global ocean uptake uncertainty. They suggest 0.7 Pg C yr\(^{-1}\) as global flux uncertainty (or approximately 53%). This number is based on a combination of error sources, including random noise in \( p_{CO_2} \) measurements, uncertainty in wind speed and the use of a constant global mean rate of change of sur-

<table>
<thead>
<tr>
<th>Study</th>
<th>Parameterization, general form: ( k = (a u_{10}^3 + b u_{10}^2 + c u_{10}) (\frac{S_c}{660})^{0.5} )</th>
<th>Global ocean carbon sink ( \text{Pg C yr}^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takahashi et al. (2009b)</td>
<td>( a = 0 ) \quad b = 0.26 \quad c = 0</td>
<td>1.42</td>
</tr>
<tr>
<td>Smethie et al. (1985)(^1)</td>
<td>( a = 0 ) \quad b = 0</td>
<td>1.33</td>
</tr>
<tr>
<td>Liss and Merlivat (1986)(^2)</td>
<td>( a = 0 ) \quad b = 0.166 \quad c = 0.133</td>
<td>0.96</td>
</tr>
<tr>
<td>Wanninkhof (1992)</td>
<td>( a = 0 ) \quad b = 0.39 \quad c = 0</td>
<td>2.13</td>
</tr>
<tr>
<td>Wanninkhof and McGillis (1999)</td>
<td>( a = 0.0283 ) \quad b = 0 \quad c = 0</td>
<td>1.78</td>
</tr>
<tr>
<td>Nightingale et al. (2000)</td>
<td>( a = 0 ) \quad b = 0.24 \quad c = 0.061</td>
<td>1.33</td>
</tr>
<tr>
<td>Ho et al. (2006)</td>
<td>( a = 0 ) \quad b = 0.266 \quad c = 0</td>
<td>1.45</td>
</tr>
<tr>
<td>Sweeney et al. (2007)</td>
<td>( a = 0 ) \quad b = 0.27 \quad c = 0</td>
<td>1.47</td>
</tr>
</tbody>
</table>

\(^1\) Adjusted constant to give global mean \( k \) of 21 cm h\(^{-1}\)

\(^2\) Original stepwise linear function was replaced by a square polynomial fit

Table 2.1  Gas transfer parameterizations from the literature.
2.4. Surface ocean pCO$_2$-based flux information

Surface ocean pCO$_2$ of 1.5 $\mu$atm yr$^{-1}$. While it also includes a contribution due to the uncertain gas transfer velocity, we decided to explicitly recalculate fluxes based on different parameterizations as mentioned above. As opposed to their scheme, this allows us to additionally account for gas transfer velocity formulations that do not follow a square dependency. Despite the huge number of $\Delta$pCO$_2$ measurements in the database, there are still regions with no observations, e.g. in the southeast Pacific (see figure 2.6). In addition to this spatial undersampling, the temporal coverage of observations varies significantly among regions. As can be seen from figure 2.6, the observational density is greater in summer than in winter. In the Southern Hemisphere winter months (June to September) less than half of the measurements were made compared to the summer months. Takahashi et al. analyze the effect on CO$_2$ fluxes of this spatiotemporal undersampling in combination with their interpolation scheme. They conclude that this could cause a bias in the flux results towards less oceanic uptake and that their global uptake of 1.4 Pg C yr$^{-1}$ may be as high as 1.6 Pg C yr$^{-1}$.

I do not apply a bias correction, but form an uncertainty component based on sampling density, in addition to the uncertainty component from the gas transfer parameterizations. For each region and month the inverse ratio of the number of samples and the typical flux variability is used to resemble that undersampling issue. The final uncertainty assigned to a region/month flux is then calculated as RMS of both uncertainty components, and lies typically between 0.2 and 0.4 Pg C yr$^{-1}$. In figure 3.1 the final ocean composite flux is shown.

In summary, the general characteristics of the resulting composite ocean flux constraint are small annual and sizeable monthly uncertainties. Annual fluxes from the ocean inversion are preferred over annual pCO$_2$-based estimates, as they are consistent with a large amount of ocean interior DIC data and do not depend on the formulation of gas exchange at the ocean’s surface. In addition, the ocean inversion provides errors as well as correlations. Figure 2.7 provides insight into the error structure for some selected regions. One can see, for example, that latitudinal correlations are much stronger than inter-basin correlations, suggesting that while the ocean inversion has difficulties to resolve CO$_2$ sinks and sources latitudinally, sinks and sources from different basins can be distinguished with great confidence. All this information is taken into account in the framework of a Bayesian synthesis inversion (section 2.5). On the other hand, ocean inverse fluxes are based on the assumption of an invariant ocean circulation, and their interannual variability is computed according to a simple model based on atmospheric CO$_2$ evolution only (see section 4.3.3 for more details). The pCO$_2$-based seasonal flux cycle is referenced to year 2000. No attempt was made to compile an interannually varying seasonal cycle, because the temporal coverage of the pCO$_2$ data is not sufficient to provide a robust basis.
Figure 2.7  Posterior fluxes, uncertainties and error correlations as estimated in the ocean inversion for selected regions. An ellipse represents the 68% confidence area for the flux estimate for each pair of regions. That is, those iso-cost lines that enclose the area where the posterior flux pair is located in with 68% probability; compare also to the projections in the conceptual drawing in figure 2.1. The cross within the ellipse gives the standard deviation for each of the fluxes. The fact that the crosses are always located in the ellipse’s interior without touching its edges is a general property of multi-variate Gaussian PDFs. The ellipse’s tilt indicates error correlation; for example, two perfectly anti-correlated flux errors would cause the ellipse to be oriented along the dashed line with slope -1. Numerical values above the ellipses give the error correlation coefficients. The top panel compares the three major ocean basins, whereas the lower panel focuses on latitudinal correlations.
2.5 Joint ocean-atmosphere inversion

In this section I describe how the atmosphere and ocean inversions are joined. The joint inversion follows the Bayesian approach, but is mathematically equivalent to the one-step Kalman approach used by Jacobson et al. (2007a), as I show below. I extend the linear synthesis method described in the context of the atmosphere inversion to the (still linear) Bayesian synthesis method by adding prior information from the ocean inversion. The joint ocean-atmosphere inversion represents the "mode 2" inversion in chapters 3 and 4. For the reader interested in more details on the implementation of the joint inversion I attached to this thesis the main Matlab routine as well as the configuration file (appendix B).

2.5.1 Bayesian synthesis inversion

Following the concepts introduced in section 2.1 to minimize model-data mismatch together with the mismatch between prior and posterior model parameters, and making use of the linearity of atmospheric transport, a cost function based on equation (2.4) can be formulated,

\[ J(x) = \frac{1}{2} (M_a x - d_a)^T C_{da}^{-1} (M_a x - d_a) + \frac{1}{2} (x - x_0)^T C_0^{-1} (x - x_0). \tag{2.27} \]

The first term represents the atmospheric model-data mismatch as considered in the atmosphere-only inversion, see equation (2.14). \( d_a \) and \( C_{da} \) are the atmospheric data vector and covariance, and \( M_a \) contains the sampled model responses at the locations of atmospheric stations. The second term represents the deviation of the model parameters \( x \) from the prior estimate \( x_0 \). The prior covariance matrix \( C_0 \) weighs the prior estimates. As described in section 2.1, the minimization of this cost function is always a compromise between fitting the atmospheric data and staying close to the prior values. The balance between data and prior constraints is defined by the data and prior uncertainties contained in the covariance matrices. In my joint inversion the posterior CO\(_2\) flux estimates from the coupled ocean inverse/pCO\(_2\)-based constraint serve as prior information. I will explain in the following how this is done in detail.

Flux estimates from the atmosphere and ocean inversions are considered statistically independent, as they are based on independent carbon measurements in the atmosphere and ocean. The estimates from the ocean approach are represented by a multivariate normal distribution \((x_o, C_o)\) for each of the 10 oceanic transport models. I chose the index \( o \) to refer to the ocean; \( x_o \) and \( C_o \) are identical to the optimized flux \( x_{opt} \) and covariance \( C_{opt} \) in equations (2.15) and (2.16). The mean vector \( x_o \) contains 132 monthly fluxes from 11 oceanic regions, averaged over the chosen inversion period. Monthly flux uncertainties are calculated by the error scheme described in section 2.4. The variances, i.e. squared uncertainties, are popu-
lated along the main diagonal of the covariance matrix $C_o$. In addition to the flux estimates and uncertainties, the ocean inversion provides information about annual error correlations between regions that are taken into account in the joint inversion, but need to be transformed into monthly correlations first. Monthly correlations are approximated by re-using the annual correlations for each month: the correlation coefficient connecting region A in January with region B in January is the same as that connecting region A in February with region B in February and so on. Monthly correlations are contained in the off-diagonal elements of $C_o$. I do not impose any "cross-month" correlation structure, so all correlation coefficients related to combinations involving two different months are set to zero. In particular, no temporal correlations within one region are considered.

The joint flux estimate and covariance is also represented by a multivariate normal distribution, $(x_j, C_j)$, and is derived by minimizing the cost function

\[
J(x) = \frac{1}{2} (M_a x - d_a)^T C_{d_a}^{-1} (M_a x - d_a) + \frac{1}{2} (H x - x_o)^T C_o^{-1} (H x - x_o),
\]

which corresponds to the general Bayesian cost function (2.27) with the ocean constraint as prior term. The matrix $H$ is necessary to map the joint estimates into the subspace of ocean inverse estimates, that is, it picks the $11 \times 12$ air-sea fluxes from the vector $x$ containing both air-sea and air-land flux parameters. Minimization of (2.28) yields (Tarantola, 2005a, chap.3)

\[
\begin{align*}
x_j &= \left(M_a^T C_{d_a}^{-1} M_a + H^T C_o^{-1} H\right)^{-1} \left(M_a^T C_{d_a}^{-1} d_a + H^T C_o^{-1} x_o\right) \\
C_j &= \left(M_a^T C_{d_a}^{-1} M_a + H^T C_o^{-1} H\right)^{-1}
\end{align*}
\]

for the posterior joint PDF $(x_j, C_j)$. Note that this reduces to the atmosphere-only solution (2.15, 2.16) if the prior terms are omitted.

As the inversion is designed to estimate monthly fluxes, the uncertainty of their annual mean is only indirectly constrained based on the monthly flux average and the error correlation structure. In order to constrain the annual mean oceanic fluxes directly by the oceanic inverse results, they are formally incorporated as additional observations. That is, both the observational vector $d_a$ and the observational covariance matrix $C_{d_a}$ are extended to accommodate 11 additional elements, representing annual flux and uncertainty estimates from the ocean inversion. The transport matrix $M_a$ is accordingly extended by 11 rows, each forming the annual mean flux from the monthly elements in the flux vector $x_j$. This allows to directly propagate the annual constraints from the ocean inversion into the joint inversion without affecting monthly fluxes. Figure 2.8 illustrates where the individual constraints enter the inversion system.

Another approach to incorporate information from an ocean inversion was used up by Jacobson et al. (2007a). They augmented an annual atmosphere inversion with results from the
2.5. Joint ocean-atmosphere inversion

$$J(x) = \frac{1}{2} (M_a x - d_a)^T C^{-1}_{d_a} (M_a x - d_a) + \frac{1}{2} (H x - x_o)^T C^{-1}_o (H x - x_o)$$

Data term
model-data mismatch

Prior term
flux update

- Ocean interior DIC
- Atmospheric CO$_2$
- Ocean surface pCO$_2$

Figure 2.8  Illustration of the linear Bayesian cost function (2.28). Arrows indicate where (i.e. in which cost function terms) the data streams from the atmosphere and ocean enter the inversion system. The data streams are atmospheric CO$_2$, ocean interior DIC and surface ocean pCO$_2$.

Ocean inversion using the Kalman filter technique. Their approach is mathematically equivalent to mine if, as in this case, the full error correlation structure in $C_o$ is propagated through the minimization process. A major difference is that my joint inversion resolves fluxes on a monthly cyclostationary timescale. Yet, for the mathematical formulation this does not make a difference, since it just increases the number of parameters contained in the model state vector $x_o$ by a factor of 12 (the number of months per year). I show the equivalence of both methods in the following paragraphs.

For their atmosphere inversion Jacobson et al. did not use any prior information, i.e. they minimized the same cost function (2.14) as I do in my atmosphere inversion. Posterior estimates from the atmosphere inversion are represented by a multivariate normal distribution $(x_a, C_a)$ with mean (optimized) flux vector $x_a$ and covariance matrix $C_a$. Jacobson et al. inverted the ocean data separately, before they constructed a joint estimate $(x_j, C_j)$ by updating the atmospheric posterior estimates $(x_a, C_a)$ with the posterior PDF $(x_o, C_o)$ from the ocean inversion:

$$x_j = x_a + K(x_o - Hx_a),$$

$$C_j = (I - KH)C_a.$$  \hspace{1cm} (2.31)  \hspace{1cm} (2.32)

Here $K$ is "optimal gain matrix" that is determined by sequential state estimation methods (Kalman, 1960; Maybeck, 1979),

$$K = C_a H^T (HC_a H^T + C_o)^{-1}.$$  \hspace{1cm} (2.33)

The matrix $H$ again extracts just the air-sea fluxes from the flux vector $x_a$. Using matrix identities developed in (Tarantola, 2005a, chap.6.30), $K$ can alternatively be written as

$$K = (C_a^{-1} + H^T C_o^{-1} H)^{-1} H^T C_o^{-1}.$$  \hspace{1cm} (2.34)
Chapter 2. Inverse method

Now I show that minimizing first (2.14) followed by forming a joint estimate using (2.31, 2.32) is equivalent to forming a joint estimate by minimizing (2.28).

The solution to the Bayesian minimization problem is given by (2.29, 2.30), whereas the minimization of the atmosphere-only cost function (2.14) yields (2.15, 2.16), or in the notation with atmosphere index $a$:

$$x_a = \left( M_a^T C_{da}^{-1} M_a \right)^{-1} M_a^T C_{da}^{-1} d_a,$$

(2.35)

$$C_a = \left( M_a^T C_{da}^{-1} M_a \right)^{-1}.$$

(2.36)

With these expressions and by use of (2.34), equations (2.29, 2.30) can be rewritten to yield

$$x_j = \left( C_a^{-1} + H^T C_o^{-1} H \right)^{-1} \left( C_a^{-1} x_a + H^T C_o^{-1} x_o \right),$$

(2.37)

$$C_j = \left( C_a^{-1} + H^T C_o^{-1} H \right)^{-1}.$$

(2.38)

Using the identity

$$(I - KH)C_a = \left( I - \left( H^T C_o^{-1} H + C_a^{-1} \right)^{-1} H^T C_o^{-1} H \right) C_a$$

$$= C_a - \left( C_a^{-1} + H^T C_o^{-1} H \right)^{-1} H^T C_o^{-1} H C_a$$

$$= \left( C_a^{-1} + H^T C_o^{-1} H \right)^{-1} \left[ (C_a^{-1} + H^T C_o^{-1} H) C_a - H^T C_o^{-1} H C_a \right]$$

(2.39)

$$= \left( C_a^{-1} + H^T C_o^{-1} H \right)^{-1} I,$$

where $I$ denotes the identity matrix, leads to

$$x_j = (I - KH)x_a + K x_o,$$

(2.40)

$$C_j = (I - KH)C_a,$$

(2.41)

which is identical to (2.31, 2.32). This makes the Kalman filter (in the one-step setup used by Jacobson et al. (2007a)) equivalent to the linear Bayesian synthesis method that I use for my joint inversion (if full ocean covariance matrix is maintained). Hence, my results can directly be compared to those of Jacobson et al., though only on an annual basis.

The inclusion of a prior constraint changes the calculation of the reduced $\chi^2$ diagnostic to

$$\chi^2 = \frac{J_{\text{min}}}{N_{\text{obs}}} = \frac{1}{N_{\text{obs}}} \left( \sum_{n=1}^{N_{\text{obs}}} \frac{(M_a x_j - d_a)^2}{\sigma_n} + \sum_{r=1}^{N_{\text{fluxes}}} \frac{(x_j - x_o)^2}{\sigma_{o,r}} \right),$$

(2.42)

where $J_{\text{min}} = J(x_j)$ is the minimum of the Bayesian cost function. $N_{\text{obs}}$ and $N_{\text{fluxes}}$ are the number of atmospheric CO$_2$ observations (not including the additional equations from the global growth and annual ocean inverse constraints) and flux parameters, respectively.
As for the atmosphere-only inversion, I apply an iterative procedure to obtain a model mean \( \chi^2 \) close to 1, achieved by adjusting the atmospheric data uncertainty. “Model-mean” now refers to the mean of the whole suite of 10 oceanic and 11 (or 12 for the 1992-1996 inversion) atmospheric transport models (table 4.1). Joint flux estimates are computed for each of the atmospheric transport models, which are represented by different transport matrices \( M_a \), while using ocean prior flux and error estimates from each of the 10 ocean inversions (ocean models also listed in table 4.1). This gives 110 (or 120) joint flux vectors and covariance matrices for each network and inversion period. The spread among this ensemble (represented as one standard deviation around the ensemble mean) is added to the mean covariance matrix by RMS, resulting in a "total" covariance matrix. As for the mode 1 atmosphere inversion, flux uncertainties reported in this thesis always refer to that total estimate.

### 2.5.2 Inclusion of further constraints

All inversions were generally performed using response fields from all atmospheric transport models. However, for my "control" inversions I selected a subset of three models according to the recommendations of Stephens et al. (2007) (cf. chapter 1). These control inversions correspond to inversion "mode 4" for the 1992-1996 period (chapter 3), respectively inversion "mode 3" for the decadal 1980-2008 periods (chapter 4). In figure 2.9 the effect of different inversion modes on the estimation of the tropical land source and northern land sink is illustrated. The illustration represents results for the 1990-1999 decadal inversion in chapter 4 for the common network. The atmosphere-only inversion (mode 1, gray ellipses) estimates large fluxes with very large uncertainties on the order of \( \pm 2 \) and \( \pm 5 \) Pg C yr\(^{-1}\) for the northern land sink and tropical source, respectively. The joint ocean-atmosphere inversion (mode 2) is shown in blue colors; posterior flux uncertainty is greatly reduced. Note that the depicted regions are located on land, that is, the ocean inversion has no direct influence on the flux estimation there; the fact that the uncertainties are nevertheless reduced so much shows how much information is contained in the ocean prior and propagated to the land via atmospheric transport. The 1990-1999 control inversion (mode 3 in chapter 4) estimate is shown in green. The green ellipse encompasses the Stephens model subset result of the blue ellipses. Mode 3 exhibits further uncertainty reduction due to the model subset's less diverse ventilation schemes (cf. discussion in chapter 4).

The model subset consists of the GISS, JMA-CDTM and TM3 models, which were identified by Stephens et al. (2007) to most closely reproduce observed vertical gradients of CO\(_2\) in the atmosphere (see table 4.1 for a list of models). Validation of models against observations in the higher troposphere is a very useful method to quantify model skills, as it is an independent test for the models, which are usually optimized to measurements in the planetary boundary layer only. In their study, Stephens et al. compared observed vertical CO\(_2\) gradients at 12
stations around the globe with model predictions. Sinks and sources of CO$_2$ were prescribed in each model according to the T3L2 set of fluxes from the same model, and then transported and sampled at the same locations as the observations. The above three models turned out to most closely reproduce the observed profiles on an annual basis. However, the models’ success in reproducing annual mean vertical gradients does not imply they have superior transport skill compared to the other models. In fact, for both the boreal winter and summer seasons, these models show vertical CO$_2$ gradients that are smaller than observed, and for each season at least two of the other models perform better than any of them (supplemental material of Stephens et al. (2007)). Because of the opposing sign of vertical gradients in summer and winter, the effects of a too strong ventilation in summer and winter compensate for each other, which is the reason for those models to reproduce annual gradients accurately.

Additional constraints for some land fluxes were available for the 1992-1996 inversion period, which I included as mode 3 in chapter 3, before restricting the atmospheric transport models in mode 4. As they are used only for that period, I describe the land constraints in section 3.3.3.
2.6 Pseudo-inversion of remotely sensed atmospheric CO$_2$ column data

An additional constraint for a joint inversion could be atmospheric column integrated (also called "total column") or column mean CO$_2$ data retrieved by spaceborne instruments attached to satellites. Such data would have the great advantage of global coverage as well as high spatial (sensor footprints roughly 10 km x 10 km) and temporal (complete global coverage roughly every 3 days) resolution.

2.6.1 Retrieval of CO$_2$ column data

There are several potential techniques (see Breon and Ciais (2010) for a comprehensive overview) for retrieving information on atmospheric CO$_2$ column concentrations from satellites:

- **Thermal infrared sounding:** The radiance emitted by the Earth's surface and the atmosphere are measured in the thermal infrared bands (wavelength longer than 4 $\mu$m) of CO$_2$. The technique requires that the surface properties, temperature profile and emissions by other gases (in particular H$_2$O) are known, because the atmosphere's composition and temperature determine the intensity of the measurable thermal infrared signal. If all these properties are known the remaining contribution of CO$_2$ can be isolated and the total column amount of CO$_2$ reconstructed. Observational coverage is limited by cloud, but the instrument works during the day as well as during the night. An example of such an instrument is AIRS (table 2.2).

- **Solar spectroscopy in the near infrared:** The idea is to measure sunlight reflected from the Earth’s surface at wavelengths in the near infrared (wavelengths shorter than 3 $\mu$m, but too long to be visible as red color), specifically in the two major absorption lines of CO$_2$ around 1.6 and 2 $\mu$m. A major advantage of this method is that it is relatively insensitive to the temperature and H$_2$O profiles. But it works only during the day when the sun is shining. Examples are the SCIAMACHY and GOSAT platforms.

- **Active sensing:** This technique is similar to the near infrared spectroscopy, replacing the sun by an artificial light source such as a Lidar (light detection and ranging). This removes the daylight constraint from the passive systems. Another potential advantage is that vertical CO$_2$ profiles could be obtained by additionally measuring the travel time of the emitted light, which allows to derive the scattering height in the atmosphere. A major difficulty with active systems is the energy supply for the Lidar device.
In general, none of the currently envisioned remote sensing techniques allows the retrieval of a vertical profile of atmospheric CO$_2$. Rather, a weighted mean column concentration is retrieved, with the vertical weighting function depending on the design of the instrument. The optimal weighting function would peak in the PBL (planetary boundary layer), i.e. the lower atmosphere, because CO$_2$ gradients in the PBL respond much more sensitive to surface sources and sinks than gradients in higher atmospheric layers. Instruments that operate in the thermal infrared have a weighting function that peaks much higher in the atmosphere than those operating in the near infrared, owing to the specific pressure dependent atmospheric transmittance in the different wavelength regimes (see Breon and Ciais (2010) for a more detailed theoretical background).

<table>
<thead>
<tr>
<th>Instrument/Platform</th>
<th>Method &amp; Status</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIRS and IASI (NASA and ESA)</td>
<td>-Thermal infrared (IR) -AIRS on AQUA -IASI on MetOp</td>
<td>-Temperature and water vapor for weather forecast -No sensitivity for the PBL -Accuracy too low to retrieve CO$_2$ column inventory -Seasonal CO$_2$ cycle of upper troposphere resolvable</td>
</tr>
<tr>
<td>SCIAMACHY (ESA)</td>
<td>-Near IR at 1.6 $\mu$m -on Envisat since 2002</td>
<td>-Moderately sensitive to CO$_2$ in PBL -Proof of concept for monitoring CO$_2$ and CH$_4$ -Too inaccurate to improve flux knowledge</td>
</tr>
<tr>
<td>GOSAT (JAXA and NIES)</td>
<td>-near IR at 0.76 $\mu$m, 1.6 $\mu$m and 2 $\mu$m -flying since 2009</td>
<td>-Dedicated to CO$_2$ and CH$_4$ observations -Sunglint capability over oceans -Not clear yet if 0.5% or 1% accuracy reachable</td>
</tr>
<tr>
<td>OCO / OCO-2 (NASA)</td>
<td>-near IR as GOSAT -launched failed</td>
<td>-Accuracy objective 1 ppm (&lt;0.4%) -OCO-2 approved, scheduled for 2013</td>
</tr>
<tr>
<td>A-Scope and ASCENDS (ESA and NASA)</td>
<td>-active sensing -A-Scope rejected -ASCENDS in airborne test phase</td>
<td>-ASCENDS airborne version achieved 0.5% accuracy -Active concept very demanding, accuracy not necessarily better than passive missions -No launch expected before 2016</td>
</tr>
</tbody>
</table>

Table 2.2 Overview on selected spaceborne instruments for CO$_2$ column retrieval.

The inclusion of satellite-derived column data in my joint inversion was not possible during the time of my PhD project, because until now there are no such data products available with sufficient precision (Breon and Ciais, 2010). The first instrument dedicated to CO$_2$ measurements, the Greenhouse gases Observing SATellite (GOSAT), was launched in 2009 and started to collect column data. However, the necessary accuracy of such measurements is currently estimated to be at least 1% (roughly 3-4 ppm), which has not been reached yet with GOSAT-derived measurements. Another dedicated satellite was NASA's Orbiting Carbon Observatory (OCO), which was launched in 2009 but failed to reach orbit due to rocket failure. Several earlier instruments are already in flight, but they were not specifically designed to measure CO$_2$. Nevertheless, attempts have been made to indirectly retrieve column concentrations of CO$_2$, but accuracy and biases remain serious issues. As a consequence, the assimilation of such data in atmospheric transport inversions brings little new information to...
our current knowledge about CO$_2$ surface fluxes. I therefore decided to study the potential benefits from columnar data in the framework of a pseudodata inversion (see also previous pseudo-inversions of, e.g., Rayner and O'Brian (2001); Maksyutov et al. (2003); Kadygrov et al. (2009); Chevallier et al. (2009)).

### 2.6.2 Setup of pseudo-inversion

A pseudo-inversion is based on simulated CO$_2$ column data instead of real data. The data are simulated by propagating surface fluxes through the same atmospheric transport model used later in the inversion to reconstruct the fluxes. Because of this, atmospheric transport is perfect in such an inversion and, hence, errors arise only due to measurement density and uncertainty.

For my inversion I simulated the column data by propagating the joint (mode 2) posterior fluxes from the 1992-1996 inversion through the TM3 (table 4.1) model. In a recent study, Pickett-Heaps et al. (2011) highlighted this model, as it performs better than other models to reproduce vertical gradients at two stations: Cape Grim, Tasmania, and Carr, Colorado. The modeling skill for vertical gradients is particularly important in the context of inverting column data that are a (weighted) composite from all atmospheric layers.

The modeled responses were combined cyclostationary to represent the monthly response averaged over a 5 years period. These three-dimensional CO$_2$ concentration fields were then sampled for each month in two ways:

1. In the PBL at station locations of the network underlying the joint 1992-1996 inversion of chapter 3 (see red symbols in figure 2.2). This is to allow for later comparison of the constraining power of the surface network versus the satellite retrievals.

2. At each of the 3312 pixels of the TM3 model (spatial resolution is 5°x4°), averaged over the whole atmospheric column (figure 2.10). This represents the maximal data density for inversions with this model, although real satellite measurements may have higher spatial resolution of about 100 km$^2$. On the other hand, I considered no cloud cover, which would reduce the number of usable measurements considerably (Rayner et al., 2002). Vertically averaging the data corresponds to applying a constant weighting function, that is, I assume a virtual satellite that detects CO$_2$ signals equally from all atmospheric depths. In order to assess the importance of the global coverage of data, I performed an additional inversion using column mean data only at locations of the surface stations (results in chapter 3).

Both pseudo-datasets are inverted using the same Bayesian approach as for the real surface
**Chapter 2. Inverse method**

**a)** Example of column mean CO$_2$ of GOSAT (April 2009 to January 2010)

**b)** Monthly (cyclostationary) column mean pseudodata in ppm (TM3 model, background response presubtracted)

**Figure 2.10** Distribution of column mean CO$_2$ data. **a)** preliminary retrievals from the GOSAT satellite for selected months in 2009 and 2010 in ppm, picture source: [http://www.gosat.nies.go.jp/index_e.html](http://www.gosat.nies.go.jp/index_e.html). **b)** simulated column mean CO$_2$ data from the TM3 atmospheric transport model for each month in a cyclostationary year. Shown are the concentration anomalies in ppm at each model pixel, according to equation (2.19) after presubtraction of the background responses.

Data-based joint inversion (cf. equation (2.28),

$$J(x) = \frac{1}{2} (M_{A} x - d_{A})^T C_{d_{A}}^{-1} (M_{A} x - d_{A}) + \frac{1}{2} (H x - x_0)^T C_{o}^{-1} (H x - x_0),$$

(2.43)
2.6. Pseudo-inversion of CO$_2$ column data

where the subscripts $a$ and $A$ represent the terms related to the surface and column inversions. That is, $M_A$ is the transport matrix containing the column mean responses at the model pixel scale, and the vector $d_A$ contains the simulated column mean CO$_2$ concentrations. The observational uncertainties prescribed in $C_a$ are identical to the uncertainties computed for the 1992-1996 surface inversion network (section 2.2.2). Hence, the pseudo surface inversion is identical to the 1992-1996 inversion, except that the Globalview-based concentrations are replaced by the simulated concentrations (but not the uncertainties). The global growth constraint as well as the background fields (fossil fuel emissions, NEP for terrestrial regions and pCO$_2$-based air-sea exchange for oceanic regions) are applied in the same way as for the 1992-1996 surface inversion. The column data uncertainties prescribed in $C_A$ are set to a constant value for all pixels. The constant is chosen between 0.1 and 5 ppm to reflect different accuracies for the satellite measurements.

The posterior flux estimates are obtained by minimization of $J$ as described in the previous sections. I performed the pseudo-inversions using both the atmosphere-only setup (without the prior term in equation (2.43) and the joint ocean-atmosphere setup (including the composite ocean constraint as prior term) to assess the additional benefit from column observations for both cases. As for the surface inversion, no land priors are used in the atmosphere-only setup, contrasting to previous pseudodata studies of, e.g., Rayner and O'Brian (2001); Maksyutov et al. (2003).
Chapter 3

A multiple-constraint inversion for the estimation of seasonal carbon sources and sinks

3.1 Abstract

We have estimated global surface fluxes of carbon dioxide using an inverse approach that sequentially considers four constraints: 1) atmospheric CO$_2$, 2) ocean interior DIC (Dissolved Inorganic Carbon) and surface ocean pCO$_2$ (partial pressure of CO$_2$), 3) annual prior fluxes for selected land regions, and 4) selection of atmospheric models based on vertical transport skill. Estimated fluxes are monthly resolved, representing the monthly mean over the period 1992-1996. Their spatial resolution corresponds to the 22 Transcom regions defined over land and ocean. The ocean constraint is particularly valuable, as it does not only add prior information about air-sea fluxes to the inversion problem, but also preserves the regional variance-covariance structure from the underlying ocean interior inversion. It allows to constrain annual oceanic uptake of 1.8 Pg C yr$^{-1}$ to within 0.2 Pg C yr$^{-1}$, which implies an annual land uptake of 1.3 ($\pm$0.3) Pg C yr$^{-1}$. Furthermore, it leads to a pronounced asymmetry in the seasonal pattern of global land uptake, which was not seen in previous atmosphere-only inversions. Tropical land is consistently estimated to be a source of carbon, though the source magnitude is reduced when more and more constraints are applied. With all four constraints, the inversion suggests a tropical source of 1.1 ($\pm$0.9) Pg C yr$^{-1}$. This is comparable to global estimates of deforestation rates in tropical forests and therefore implies an annually balanced tropical land biosphere flux. This balance is not found, however, at the regional level: for the Amazonian region we find a biospheric source of 0.6 ($\pm$0.5) Pg C yr$^{-1}$, somewhat at the upper range of estimates from bottom-up methods, which on average suggest this region to be a sink for carbon. We find only a small source over aggregated Tropical and Southern
Hemisphere Land (TSL) regions, owing in large parts to the applied selection of transport models. All constraints point towards a seasonal TSL amplitude much larger than estimated by previous studies. While all constraints have a significant impact on flux estimates, the ocean constraint has the additional value to make land fluxes consistent with measurements taken in the ocean.

3.2 Introduction

In the 1990s, 64 Pg C were emitted to the atmosphere through fossil fuel burning and cement manufacture (Boden et al., 2010). A portion of 33 Pg C remained in the atmosphere, corresponding to 52% of the emissions. This caused an increase in \( \text{CO}_2 \) concentration from 353 to 367 ppm, calculated as average of Mauna Loa and South Pole time series data (cf. Globalview-co2 (2009)). The remaining 31 Pg C have been taken up by the ocean and the land biosphere. Their combined sink strength is known with a high degree of confidence, because of the high accuracy of both the fossil emissions estimate and the \( \text{CO}_2 \) records. However, the magnitudes of the individual land and ocean sinks are less certain, and additional regional breakdown (for example into continental-/basin-scale budgets) increases the uncertainties further.

Several methods have been applied to separate the ocean from the land sink. Some of them use the fact that oceanic uptake of carbon does not fractionate (e.g. Battle et al. (2000), \( \delta^{13}C \) method) and is independent of the uptake of other gases, such as oxygen (\( \text{O}_2/\text{N}_2 \) method, e.g. Battle et al. (2000), Plattner et al. (2002), Bopp et al. (2002)). On the other hand, the exchange of carbon and oxygen between the land biosphere and the atmosphere is strongly correlated, due to the \( \text{O}_2/\text{C} \) stoichiometry that underlies the processes of photosynthesis and respiration. The land sink is then directly determined from the biospheric imprint on these tracers. Those studies show that in the 1990s, the land must have taken up less carbon than the ocean, in particular when a possible climate-induced net outgassing of oceanic oxygen is taken into account (Plattner et al., 2002; Bopp et al., 2002). Another possibility is to use surface ocean \( p\text{CO}_2 \) data to determine the air-sea carbon flux portion on the basis of the difference to \( p\text{CO}_2 \) in the atmospheric boundary layer over the oceans (e.g. Takahashi et al. (2009b)). The terrestrial sink can then be calculated as the residual of atmospheric growth rate and air-sea flux. These studies suggest a smaller ocean sink and therefore attribute most of the combined uptake to the land biosphere.

Oceanic and atmospheric carbon sinks can also be estimated by interpreting gradients in atmospheric \( \text{CO}_2 \) concentration through a transport inversion. Inferences about surface fluxes from atmospheric measurements have a long history (e.g. Keeling (1960)) and have led to the development of sophisticated inversion methods on various spatiotemporal flux scales.
3.2. Introduction

(e.g. Bousquet et al. (1999b), Gurney et al. (2002, 2004), Roedenbeck et al. (2003a), Baker et al. (2006), Peters et al. (2007)). Every inversion method is comprised of three main components: a set of atmospheric CO₂ data, an atmospheric transport model, and a regional set of surface fluxes. The regional fluxes are chosen such that their propagation through the transport model most closely resembles the observed atmospheric CO₂ data. In a similar manner, gradients in dissolved inorganic carbon (DIC) in the interior ocean can be interpreted with an ocean circulation model to estimate sea-air fluxes of CO₂ for prescribed surface regions (e.g. Gloor et al. (2003), Mikaloff Fletcher et al. (2006, 2007), Gruber et al. (2009)). Results from ocean inversions usually suggest a larger ocean sink. Gruber et al. (2009) found -1.7 Pg C yr⁻¹ (sign convention assigns negative fluxes as out of the atmosphere) for long periods of the 1990s. When combined with the above growth estimate, this leaves -1.4 Pg C yr⁻¹ as land sink. The atmospheric inversion of Gurney et al. (2004), on the other hand, estimates a smaller ocean sink of -1.3 Pg C yr⁻¹, and therefore attribute -1.8 Pg C yr⁻¹ to the land.

This study is motivated by the discrepancy of previous results and aims to reconcile flux estimates by a joint inversion of atmospheric CO₂ and ocean interior DIC data. Flux results represent monthly averages for the 1992-1996 period for 11 ocean and 11 land regions, corresponding to the Transcom regions set (e.g. Gurney et al. (2002)). Ocean inversion results always represent longterm mean fluxes due to the long timescales of ocean circulation, as opposed to the rapid atmospheric circulation that makes atmospheric inversions capable of resolving fluxes on monthly timescales. We use additional information about seasonal variability from a recent dataset of surface ocean pCO₂ (Takahashi et al., 2009b) to form the aforementioned monthly average flux distribution. This approach overcomes the issue of potential misallocation of seasonal flux variability from some terrestrial regions to adjacent ocean regions, which shows up in seasonal atmosphere-only inversions (Gurney et al., 2004). While similar to the joint approach of Jacobson et al. (2007a), our method is not restricted to annual mean fluxes, but resolves their seasonal cycle. This not only allows to get new insights into the seasonal carbon cycle over land, but also avoids the need to make any assumptions about a rectifier effect due to the seasonal coupling between atmospheric circulation and carbon exchange with the land biosphere. We achieve this by coupling the ocean inversion, augmented with seasonal flux information based on the surface pCO₂ network, to a seasonal atmospheric inversion. Hence, the resulting regional fluxes are formally consistent with atmospheric CO₂, ocean interior DIC and ocean surface pCO₂ observations.

When Jacobson et al. (2007a,b) did their joint inversion to reconcile ocean-based with atmosphere-based carbon fluxes, they found very strong, compensating fluxes in certain regions. Tropical land appeared as a huge source of more than 4 Pg C yr⁻¹, and the Southern Hemisphere land as a large sink of almost -2.5 Pg C yr⁻¹. This would imply that tropical forests are a net source of carbon to the atmosphere, even when global deforestation rates are taken into account. On the other hand, bottom-up studies based on forest inventories and eddy
covariance flux towers, indicate undisturbed tropical forests to be a net carbon sink, or nearly neutral, over the year (e.g. Phillips et al. (1998), Phillips et al. (2009), Chave et al. (2008), Lewis et al. (2009)). In this study we revisit this apparent inconsistency for the tropical land between the land-based tropical constraint and the constraint from the ocean inversion.

From the observed latitudinal gradient of atmospheric CO\textsubscript{2}, inverse models commonly infer a large sink of CO\textsubscript{2} over temperate land areas in the northern hemisphere (e.g. Tans et al. (1990), Peters et al. (2007), Pacala et al. (2007)), while tropical land regions are either nearly balanced or rather weak carbon sinks. In their recent work, Stephens et al. (2007) challenge this view through an analysis of 12 atmospheric transport models that have been used for atmospheric inversions before (Transcom models, see e.g. Gurney et al. (2002)). Their model skill analysis is based on the models’ ability to reproduce observed vertical CO\textsubscript{2} gradients, which is known to be very important for surface flux estimation (Bousquet et al., 1999b). They conclude that those models with the most accurate annual mean vertical gradient are characterized by weak northern and strong tropical land carbon uptake, as opposed to the common view. For this study we confined our transport models to the subset recommended by Stephens et al. (2007) to see if their conclusions also hold in a joint ocean-atmosphere inversion.

In this study numerous constraints on the carbon cycle are considered: ocean DIC and pCO\textsubscript{2}, atmospheric CO\textsubscript{2}, bottom-up estimates for large-scale land regions, and model selection according to Stephens et al. (2007). Two aims are pursued: 1) to characterize the information content of each constraint and to identify areas of contradiction, and 2) to combine them in a joint inversion to estimate a set of regional and seasonal fluxes that is formally consistent with all of them. We begin in the next section with a description of the joint inversion and all its ingredients, i.e. data streams and constraining elements. We then turn to the regional flux results and the discussion about the role and strength of each constraint, as well as the impact on key regions, such as the tropics and temperate Northern Hemisphere land.

### 3.3 Methods

The joint inversion interprets carbon observations from the atmosphere and ocean to estimate exchange fluxes of CO\textsubscript{2} between the atmosphere and both the terrestrial biosphere and the surface ocean. To link observations and fluxes, transport models are used to simulate the distribution of CO\textsubscript{2} in the atmosphere and DIC in the ocean. The global surface is partitioned into 11 terrestrial and 11 oceanic regions, for each of which a steady-state distribution of CO\textsubscript{2}/DIC is simulated in response to a unit dye tracer flux from that region. These regional footprints, or Green’s functions, are then combined in such a way that, in a linear least squares sense (e.g. Tarantola (2005b)), their weighted sum most closely matches the
3.3. Methods

We use a suite of 10 oceanic and 12 atmospheric transport models to express transport uncertainty and to assess its impact on the inverse results. In the following, the ingredients of the joint inversion framework are elucidated in detail, which are 1) the atmospheric inversion, 2) the constraint from ocean DIC and pCO$_2$ observations, 3) additional terrestrial constraints and 4) the construction of the joint flux estimate.

3.3.1 Unregularized inversion of atmospheric CO$_2$

The atmospheric inversion combines monthly atmospheric CO$_2$ observations with information about transport from 12 different atmospheric transport models to estimate exchange fluxes between the atmosphere and both land and ocean regions. For the inversion the same transport models are used as in the Transcom3, Level 2 (T3L2) study (Gurney et al., 2004) and the NOAA Globalview CO$_2$ data set (Globalview-co2, 2009). The setup generally follows the T3L2 control inversion described by Gurney et al. (2004), but with some important deviations. A major difference is that we do not use any regularization techniques, i.e. we do not augment the inverse problem with prior information about fluxes and uncertainty structure. Further differences relate to the preparation of the observational network, the assignment of observational errors and the global fossil fuel emission estimates (see section 2.2 for details).

For the atmospheric inversion, the global surface ocean is partitioned into 11 regions. Likewise, the global land surface is divided into 11 terrestrial regions, guided by patterns of vegetation type (figure 2.2). CO$_2$ fluxes are estimated for all 22 regions and 12 months representing the stationary seasonal cycle of the inversion period, i.e. 1992-1996. For each region and month a dye tracer Green’s function is simulated, representing the steady-state atmospheric dye concentration in response to a pulsed dye flux from that region and month. These concentration fields are then sampled each month at locations of observations. During the inversion process, these region/month responses are combined to best match the observed CO$_2$ concentrations. The spread among flux estimates arising from the use of 12 transport models is interpreted as uncertainty in transport. This “between-model” uncertainty is added to the main diagonal of the internal “within-model” covariance as squared sum. The resulting “total” covariance is what will be reported in this paper.

Atmospheric CO$_2$ data to be inverted stem from the recent 2009 version of the NOAA Globalview CO$_2$ data set (Globalview-co2, 2009). We invert monthly observations averaged over 1992-1996 from 85 stations. The choice of stations is based on criteria, which ensure that the selected stations provide measurements with sufficient total (over 1992-1996) and monthly coverage (see section 2.2.2 for details). We explicitly exclude observations from the WITN surface and Darwin stations owing to concerns about data quality and local representativeness (Gurney et al., 2003; Law et al., 2003). Data uncertainty is assigned based on the
measurement density as well, but additionally depends on the monthly residual standard deviation (RSD) of the direct measurements around the smoothed data series of Globalview. Furthermore, the data uncertainty is adjusted by a global constant to achieve a model-mean $\chi^2$ close to 1.

While the number of stations listed in the Globalview-co2 (2009) data set has doubled compared to the year 2000 version used by T3L2, our station network is very similar to theirs, as most new stations started recording only recently (e.g. after 1996) and are therefore not included. Additionally, not all of the new stations covering 1992-1996 fulfill the selection criteria described above. Some of the rejected stations, however, have an influence on the computation of the marine boundary layer reference data, which underlies the smoothing procedure of Globalview-co2 (2009). This leads to different residual concentrations, and hence to different RSD values, which underlie the data uncertainty estimation scheme. The slight network change as well as changes in the Globalview dataset need to be kept in mind when comparing our results with the Transcom inversions.

The modeled response to the release of fossil fuel CO$_2$ is presubtracted from the atmospheric CO$_2$ observations, so that the inversion is actually constrained by CO$_2$ concentration anomalies (see section 2.2.3 for details). Fossil CO$_2$ emissions sum up to 6.3 Pg C yr$^{-1}$ globally for the 1992-1996 period, which is a slightly higher estimate than the value of 6.1 Pg C yr$^{-1}$ used by T3L2. As for the T3L2 study, two more background fields are presubtracted from the observations: a seasonal ecosystem model-based estimate of NEP (annual mean is zero) and a surface ocean pCO$_2$-based estimate of air-sea CO$_2$ exchange. These background fields do not represent a prior constraint in the Bayesian sense.

In addition to the observational data, the global CO$_2$ uptake by the oceans and the terrestrial biosphere constrains the inversion. It is derived as the difference between global fossil fuel emissions and observed atmospheric CO$_2$ growth rate (see section 2.2.4 for details). For the 1992-1996 period we derive an atmospheric growth rate of 3.28 ($\pm$0.16) Pg C yr$^{-1}$, as opposed to 3.27 ($\pm$0.07) Pg C yr$^{-1}$ of T3L2. As a result from similar growth rates but different fossil fuel emission estimates, global uptake is increased by about 10% in our inversion, i.e. 3.12 Pg C yr$^{-1}$ compared to 2.83 Pg C yr$^{-1}$ for the Transcom setup.

The atmospheric inversion is set up to estimate the 1992-1996 mean monthly CO$_2$ fluxes from 11 oceanic and 11 terrestrial regions around the globe. Fluxes represent net CO$_2$ exchange excluding emissions due to fossil fuel burning, cement manufacture and gas flaring, which are prescribed in the inversion. The global CO$_2$ flux is constrained by a data-based estimate of the global atmospheric CO$_2$ growth rate. No additional constraints are applied, in particular no prior information is used to regularize fluxes.
3.3.2 Ocean constraints

A combined data stream of surface ocean pCO$_2$ and ocean interior DIC underlies the derivation of the mean seasonal cycle over all oceanic regions. These estimates are included later as prior constraints in mode 2 of the joint inversion (cf. 3.3.5). Ocean interior DIC data are interpreted by an inversion and yield long-term mean fluxes and covariances, while surface ocean pCO$_2$ data are used to extract the mean seasonal flux pattern.

Ocean interior DIC inversion

The ocean inversion uses the Green’s function approach, similar to the atmosphere inversion. Ocean interior data to be inverted are available to a large extent in the GLODAP version 1.0 database (Key et al., 2004), augmented with some data collected during historical cruises in the Atlantic (Mikaloff Fletcher et al., 2006). The total number of available observations exceeds 68,000, which is one of the major differences compared to the atmosphere inversion, which suffers from data scarcity as it is constrained by 85 stations with monthly observations, i.e. only about 1,000 data points. Ocean interior observations represent carbon concentrations averaged over the time scales of ocean circulation, which is why air-sea CO$_2$ fluxes estimated by the ocean inverse method constitute long-term mean fluxes and do not provide interannual information, nor do they resolve the seasonal cycle. For more details on the method see section 2.3.

The ocean inversion does not estimate regional fluxes of CO$_2$ based on direct observations of DIC, but separately inverts data-based inventories of anthropogenic carbon and natural, preindustrial carbon. The anthropogenic portion is separated from the measured DIC using the $\Delta C^*$ method of Gruber et al. (1996). For a more detailed explanation of the method we refer to section 2.3.1.

For the Green’s function approach, a set of regional footprints is needed for both tracers. Each footprint essentially is a simulated three-dimensional concentration field of a dye tracer, resulting from a continuous injection of this tracer into one of the 30 regions. These simulations were performed using 10 different OGCMs within the scope of the Ocean Inversion Project (OIP). The models distinguish themselves by different vertical and along-isopycnal mixing schemes as well as variable water mass transformation rates in different ocean areas. Hence, the participation of several models allows us to reflect the uncertainty in transport simulation in the inversion. The spread of the inverse results among models serves as an estimate of this uncertainty. However, all models may have similar structural deficiencies, such as coarse resolution, similar parameterization of processes on subgrid scales, preventing a complete reflection of transport uncertainty. To compare with the data, the three-dimensional footprints are then sampled at locations where we have data points. CO$_2$ fluxes are aggregated to the 11 oceanic Transcom regions used in the atmosphere inversion.

A contemporary flux estimate is formed by summing preindustrial and anthropogenic flux components (section 2.3). Because the joint inversion (section 3.3.5) is designed to estimate fluxes representing the 1992-1996 period, the anthropogenic component is referenced to the midpoint of this period before summing (section 2.3). The riverine carbon loop is accounted for as described in section 2.3).

The ocean inversion is over-determined due to the large amount of available data. This leads to very small posterior flux uncertainties, typically around 0.01 Pg C yr$^{-1}$, less than 3% of the flux, for all transport models. Compared to this internal uncertainty, the spread among models is dominant in all regions. When combined as squared sum, total uncertainty is on average 0.06 Pg C yr$^{-1}$ for the 11 regions (figure 2.7).

**Surface ocean $\Delta$pCO$_2$**

The long timescales of ocean circulation together with the time-integrated character of ocean interior DIC measurements prevent the inverse estimation of air-sea CO$_2$ fluxes on monthly timescales, but for our monthly cyclostationary joint inversion we need monthly mean ocean priors. The necessary information on seasonal flux variability is extracted from the recent pCO$_2$ climatology of Takahashi et al. (2009a), and subsequently mapped on the longterm mean ocean inverse flux estimates. Monthly uncertainties are assigned to the pCO$_2$-based flux estimates, based mainly on the uncertain gas exchange rate and the spatiotemporal density of pCO$_2$ measurements. For more details on the uncertainty scheme as well as on how the ocean inverse and pCO$_2$-based flux estimates are combined, see section 2.4.

### 3.3.3 Land constraints

So far, no *a priori* information about terrestrial fluxes has entered the inversion system. This has the advantage that it isolates the information contained in the atmospheric and oceanic carbon data streams and preserves the error-correlation structure of the results. However, to abandon any land prior information also suggests to the inversion that nothing is known about CO$_2$ fluxes over land and that any set of land flux estimates is equally likely. In fact, we certainly have founded knowledge about these fluxes in some regions, which should be reflected in the inversion. At the same time, we would like to keep the character of the inversion to be mainly driven by atmospheric and oceanic carbon concentration data, and not so much by terrestrial priors.

We chose to use rather weak priors for only four selected large-scale land areas, and representing only the annual mean fluxes over the inversion period. The priors cover land areas in the Tropics and the Northern Hemisphere. In table 3.1 the prior fluxes, uncertainties, and
respective references are listed.

<table>
<thead>
<tr>
<th>Land region</th>
<th>Land-to-air annual flux (Pg C yr(^{-1}))</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td>(-0.5 \pm 0.3)</td>
<td>Pacala et al. (2007)</td>
</tr>
<tr>
<td>Eurasia</td>
<td>(-0.5 \pm 0.4)</td>
<td>Goodale et al. (2002); Sarmiento et al. (2010)</td>
</tr>
<tr>
<td>Tropical America</td>
<td>(-0.1 \pm 0.7)</td>
<td>Phillips et al. (2009)</td>
</tr>
<tr>
<td>Tropical Africa and Asia</td>
<td>0.0 (\pm 0.8)</td>
<td>Chave et al. (2008); Sarmiento et al. (2010)</td>
</tr>
</tbody>
</table>

Table 3.1 Annual carbon fluxes for aggregated land regions used to constrain the joint inversion.

As opposed to other inversion studies, such as T3L2, no direct assumptions about the terrestrial seasonal flux cycle are made. For tropical regions, where only a small seasonal flux amplitude is expected, we added a smoothing constraint for month-to-month flux variations. The constraint was implemented as Gaussian with zero mean and 5 Pg C yr\(^{-1}\) standard deviation. This has a smoothing effect on monthly flux estimates for those regions. As this would also tend to decrease monthly uncertainty estimates, uncertainties are always reported based on the non-smoothed inversion.

### 3.3.4 Selection of atmospheric transport models

For the final mode of our joint inversion setup, we restricted the atmospheric transport models to three, according to the recommendations of Stephens et al. (2007). The three models are highlighted in table 4.1 in represent those that showed most skill in reproducing annual mean vertical gradients in tropospheric CO\(_2\) (see also section 2.5.2). There is ongoing debate whether this subset of models is to be preferred, because on a seasonal basis they actually perform worse than most other models.

### 3.3.5 Joint Inversion

The joint inversion is set up in four steps (modes) to sequentially take into account the four constraints:

1. atmospheric CO\(_2\) concentration data,

2. ocean interior DIC and surface pCO$_2$ data,
3. inventory-based land fluxes for large-scale areas, and
4. atmospheric model selection based on vertical transport skill.

The inversion setup with all constraints is what will be called the "control" inversion. The control inversion is consistent with all available data streams and its results therefore represent our best guess. The joint inversion is generally performed iteratively to achieve a mean $\chi^2$ value of 1.

Fluxes from the unregularized atmosphere inversion are augmented by prior information about air-sea fluxes from the ocean inversion following the Bayesian synthesis method (Tarantola, 2005b; Enting, 2002), see chapter 2 and especially section 2.5.

The land constraints introduce only information about annual mean fluxes and uncertainties, and do not provide monthly prior fluxes. Unlike the ocean prior they are not incorporated into the cost function as an additional prior term. Instead, they enter the inversion formally as additional observations, similar to the annual constraints from the ocean inversion (see section 2.5 for more details). The observational vector is expanded by the four annual mean flux estimates and the observational covariance matrix is extended by four rows carrying the annual mean uncertainties on their main diagonal. The transport matrix is extended accordingly.

The joint inversion is performed for each of the 10 oceanic transport models in combination with each of the 12 atmospheric transport models. For all of these 120 model pairs a joint flux vector is obtained. For the control case the number of atmospheric transport models is reduced to three (the Stephens subset), which leaves joint results for 30 model pairs. From these the mean flux vector is calculated, representing the control result. To estimate the uncertainty, two contributions are considered: 1) the mean covariance matrix from the 30 model pairs and 2) the covariance of the mean fluxes among the 30 model pairs. The latter reflects the spread of flux results among models and serves as estimator for transport uncertainty. Despite the different nature of these two contributions, they are combined as squared sum for simplicity. The resulting "total" uncertainty is what is reported throughout this paper.

Inversion of simulated atmospheric CO$_2$ column observations

As described in section 2.6, we also performed pseudo-inversions of atmospheric CO$_2$ column data to assess the information content provided by such data and to compare it to the information content of the atmospheric CO$_2$ surface network. Unfortunately, because no real satellite-derived CO$_2$ column measurements are available to date, we could not incorporate them as an additional data stream in the joint ocean-atmosphere inversion.
The pseudodata were simulated with the TM3 transport model, which was also used to infer the fluxes, thus eliminating transport error. To simulate the data we propagated our mode 2 flux results through the TM3 model, that is, fluxes that are already optimized with regard to the atmospheric surface CO\(_2\) data and the oceanic carbon data. Our main focus was to determine a minimum accuracy such column measurements must have in order to compete with the existing network of surface CO\(_2\) observations. We performed first an inversion using only the pseudodata, and then using both the pseudodata and the ocean constraint. For each inversion we varied the pseudodata uncertainty between 0.1 and 5 ppm. As a measure for the inversion's ability to infer the fluxes we used the median of the annual posterior flux uncertainties.

### 3.4 Results and Discussion

In this section we present the results of our control inversion and compare them to flux estimates from previous studies. These include the seasonal Transcom 3 Level 2 inversion of Gurney et al. (2004) (henceforth "T3L2"), which differs from our approach in two main points: 1) it is a purely atmospheric inversion, and 2) they applied explicit monthly prior fluxes for all land regions. Another study is the annual joint atmosphere-ocean inversion of Jacobson et al. (2007a,b) (henceforth "Ja07"), who incorporated information from an ocean interior inversion in a similar manner as we did, but remained limited to the estimation of annual mean fluxes. Both studies represent the same 1992-1996 period, making our results directly comparable to theirs.

In addition to comparing our control results to other studies, we put particular emphasis on the evolution of flux estimates when more and more constraints are added to the inversion setup. As explained in the previous section these include 1) atmospheric CO\(_2\) data, 2) ocean interior DIC and surface ocean pCO\(_2\) data, 3) annual prior information for selected large-scale tropical and northern land regions as well as a month-to-month regularization of tropical land fluxes, and 4) transport model selection. These four steps are applied in this order, and every step includes all previous constraints as well. Our control inversion contains all constraints and therefore represents the final step, or mode 4.

The section is structured such that we first discuss sea-to-air fluxes from oceanic regions, with primary focus on differences to the T3L2 study. This is followed by a detailed discussion of land-to-air fluxes, which is guided by moving latitudinal from the northern boreal and temperate regions to the Tropics and further to the southern land regions.

Posterior uncertainties include Bayesian error estimate and the spread of results among model pairs, the latter represented by one standard deviation. Unless otherwise noted, re-
ported values combine both components by root mean square.

3.4.1 Sea-to-air-fluxes

Annual sea-to-air fluxes are well constrained in our study and the one of Ja07, due to the small flux uncertainties associated with the ocean inversion. The weighted model mean global annual sea-to-air flux is constrained to $-1.78 \pm 0.24$ Pg C yr$^{-1}$ by the ocean inversion (Mikaloff Fletcher et al., 2006, 2007; Gruber et al., 2009). Together with the well constrained global uptake, this leaves $-1.25 \pm 0.29$Pg C yr$^{-1}$ as annual mean global land-to-air flux, in agreement with the T3L2 estimate of $-1.47 \pm 0.97$ Pg C yr$^{-1}$, but with a major uncertainty reduction of 70%. Ja07 estimated a smaller terrestrial sink of $-1.12 \pm 0.23$ Pg C yr$^{-1}$, mainly driven by their smaller estimate for global uptake.

Our annual oceanic prior is the same as that of Ja07. As it represents a strong constraint for the joint inversion, our posterior estimates do not differ significantly from the prior (figure 3.1). They are also very similar to those of Ja07 in all regions (table 3.2). Considerable differences exist, however, to the T3L2 study, in particular for the tropical and SH oceans. T3L2 estimate global tropical oceans to be nearly neutral and Southern Hemisphere oceans to be a weak carbon sink of $-0.49 \pm 0.64$ Pg C yr$^{-1}$, while our control inversion estimates a tropical outgassing of $0.58 \pm 0.10$ Pg C yr$^{-1}$ and a strong carbon sink of $-1.41 \pm 0.15$ Pg C yr$^{-1}$ for Southern Hemisphere oceans.

T3L2’s neutral carbon balance in tropical oceans is a result of compensation of strong outgassing from eastern tropical Pacific with uptake of carbon in the other tropical basins (western tropical Pacific, tropical Indian and tropical Atlantic, see figure 3.1 and table 3.2). Opposed to that, the control inversion suggests an annual mean outgassing across all tropical oceans. In the Southern Hemisphere, major differences exist in the Southern Ocean and, in particular, in the south Pacific, where T3L2 estimate strong outgassing, while we find a significant uptake.

These discrepancies in annual mean fluxes from individual tropical and southern regions do not compensate when aggregated latitudinally: the probability that aggregated tropical oceans are a source of carbon (and not neutral as estimated by T3L2) is greater than 99% (based on one-sided Gaussian cumulative density function). The probability that our Southern Hemisphere ocean sink is larger than the sink estimated by T3L2, is greater than 99% as well.

The seasonal variability of the prior fluxes is less constrained than the annual mean, as it is not based on the ocean inversion, but on $\Delta pCO_2$ measurements of Takahashi et al. (2009b). The precision of monthly flux estimates is mainly limited by incomplete knowledge about gas
3.4. Results and Discussion

Table 3.2 Annual posterior fluxes for all 22 regions in the inversion as well as some additional large-scale regions. Results from all inversion modes are compared to the studies of Gurney et al. (2004) (T3L2) and Jacobson et al. (2007b) (Ja07). Bold numbers indicate 90% statistical significance for the flux sign.

Transfer rates as well as spatial coverage of the pCO₂ measurements. As a result, the addition of the land and transport constraints has an effect on monthly posterior flux estimates. However, this effect remains very small in most regions. Exceptions are the northern oceans, which appear more seasonally variable than suggested by the pCO₂ prior, and the Southern Ocean, whose seasonal amplitude appears larger. The control inversion estimates a Southern ocean outgassing of 0.3 Pg C accumulated over months July to October. Opposed to
that, T3L2 found the Southern Ocean to be a constant sink throughout the year with much smaller seasonal amplitude, which is inconsistent with pCO$_2$-based estimates. Another issue with the weak ocean prior used by T3L2 is that it allows the inversion to erroneously allocate seasonal flux variability to ocean regions, which is partially caused by seasonal flux variability over adjacent land areas.

Our joint inversion resolves this issue by strengthening the sea-to-air flux priors. However, this strengthening is not arbitrary, but follows the twofold uncertainty scheme described in the previous section: the monthly pCO$_2$-based uncertainty and the annual uncertainty obtained from the ocean inversion. Therefore, joint inverse fluxes are constrained by, and consistent with, two different data sources, corresponding to the annual and monthly time scales. These well-constrained sea-to-air fluxes propagate information to the land, where they influence regional fluxes, which can then as well be estimated in consistence with the oceanic DIC and pCO$_2$ datasets.

### 3.4.2 Northern extra-tropical land

Fluxes from Northern Hemisphere extra-tropical land (boreal and temperate) are generally better constrained by atmospheric CO$_2$ concentrations than from tropical or Southern Hemisphere land, due to the better observational coverage (figure 2.2). However, the inclusion of the ocean constraint in mode 2 of our joint inversion reduces the annual northern extra-tropical sink significantly (figure 3.3). It also drives the uncertainty below 2 Pg C yr$^{-1}$. The main effect of further inclusion of land and model constraints in modes 3 and 4 is a continued uncertainty reduction until the control estimate of $-1.8 \pm 0.4$ Pg C yr$^{-1}$ is obtained.

Annual mean control results for the boreal regions in America and Asia are in very good agreement with the studies of T3L2 and Ja07 (figure 3.2, table 3.2). They are also not sensitive as to which constraints are included: already the pure atmospheric inversion (mode 1) estimates fluxes very similar to the control (mode 4) fluxes. Seasonal variability is also very consistent between T3L2 and all modes of the joint inversion. Uncertainties decrease when more constraints are added, but not significantly. Fluxes from those regions are robust and can be determined by inversion of atmospheric CO$_2$ observations alone.

Annual uptake by northern temperate land is estimated to $-1.5$ Pg C yr$^{-1}$, smaller than the estimates of T3L2 and Ja07 ($-2.3$ Pg C yr$^{-1}$ and $-2.8$ Pg C yr$^{-1}$, respectively). The reason is mainly that the control inversion suggests Temperate Asia a weak source of carbon, while both Ja07 and T3L2 suggest a considerable sink. The switch from sink to source is caused by the inclusion of the land constraint in the inversion (see difference in annual mean flux between mode 2 and 3 in figure 3.2). In mode 3 the Eurasian land region (boreal and temperate) was constrained to $-0.5 \pm 0.4$ Pg C yr$^{-1}$. Because the boreal Asian region as well as
3.4. Results and Discussion

Figure 3.1 Black lines show the posterior results from the control (mode 4) inversion for the 11 ocean regions. According to the sign convention, a positive flux is directed into the atmosphere. Gray shaded areas represent the total posterior uncertainty as one standard deviation, which in turn includes contributions from the Bayesian "within" error as well as from the "between" error arising from transport model spread. Results are compared to the T3L2 (Transcom 3, Level 2) atmosphere-only inversion of Gurney et al. (2004) in green, for which the monthly uncertainty is not shown for clarity. In light blue (uncertainty given by the dashed line) the ocean constraint is shown, as included in joint inversion mode 2.
Figure 3.2  Control fluxes for the land regions are compared to results obtained in each mode of the joint inversion.
Europe are well-constrained by atmospheric CO$_2$, the joint inversion preferably adjusts the flux from temperate Asia to achieve matching the prior.

Over temperate Asia, CO$_2$ uptake during the growing season is greatly reduced compared to T3L2. Not only the total uptake is reduced, but also the length of the growing season. It covers the months May to July, whereas it extends until September in the T3L2 study. From our 4-step analysis we see this reduction is a consequence of the ocean constraint (pink curve in figure 3.2), which corrects the estimate from the atmosphere-only (mode 1) inversion. Further inclusion of the land as well as the transport constraints does not alter this picture much more. It is interesting to note that this correction is not due to a simple reallocation of flux from adjacent ocean regions. For example, the T3L2 strong seasonal signal in the Tropical Indian ocean is not directly propagated to the Temperate Asia region in the North, as this would result in enhanced CO$_2$ uptake there during the growing season. The same is true for the other two adjacent ocean regions, i.e. the North and Western Tropical Pacific. Hence, the reduced growing season uptake is a result of the appliance of the complete ocean prior constraint, not only few selected regions.

Compared to T3L2, our control inversion estimates a slightly smaller European sink, driven by enhanced winter respiration. This is compensated by an increased sink over temperate North America (USA plus parts of Mexico and small parts of Canada, cf. figure 2.2). The latter is a result of enhanced growing season uptake, partially offset by slightly increased winter respiration. Based on the 4-step analysis we see that the increased winter respiration in these regions is primarily a consequence of the transport model selection (mode 3 vs. mode 4). This is due to the ventilation characteristics of the selected models, as described in one of the following sections. The uncertainty of the annual mean temperate North American sink is reduced significantly only in mode 4 of our inversion series, i.e. after model selection. This is due to one particular atmospheric model, the NIRE model, which estimates an uptake of nearly -6 Pg C yr$^{-1}$. This results in a large model spread and accordingly a particularly high error estimate, despite the good observational coverage of the region.

### 3.4.3 Tropical land

We estimate a net annual pantropical land source of 1.1 ($\pm$0.9) Pg C yr$^{-1}$ in our control inversion. When a global deforestation rate of 1.3 ($\pm$0.8) Pg C yr$^{-1}$ (Sarmiento et al., 2010) is taken into account, this implies a biospheric sink of 0.2 ($\pm$1.7) Pg C yr$^{-1}$ for tropical regions. When taking a closer look at the regional level, it becomes evident that the net source is solely located in the Tropical American region, which in turn consists mainly of the Amazonian tropical rainforest. A more detailed discussion about implications for Amazonia follows later, while this section is more designated to our results for the remaining tropical regions in Asia and Africa, as well as their intercomparison and sensitivity to the inclusion of mode 1 to

Total uncertainty in PgC/yr annual CO₂ flux into the atmosphere in PgC/yr

- 5
- 4
- 3
- 2
- 1
0
1
2
3
4
5
6
7
0.3
0.5
1
2
3
5

Figure 3.3 The total uncertainty (squared sum of “within” and “between” error contributions; one standard deviation; logarithmic axis) is plotted against the annual mean flux from aggregated, latitudinal land regions. Yellow colors separate global land into Northern Hemisphere extratropical (NH) land and aggregated Tropical and Southern Hemisphere Land (TSL). Green colors further partition TSL into Tropical land and Southern Hemisphere extratropical (SH) land. Gray shaded areas represent flux sign significance regimes between 70% and 99%, i.e. a point located in such a regime represents a flux that is positive (or negative) with the indicated probability. The numbers inside the symbols give the inversion mode: number 1 represents the atmosphere-only inversion, while number 4 represents the control inversion. Black lines show the trajectory of inverse results when more and more constraints are considered.

4 constraints.

As opposed to Tropical America, the control inversion finds the North African region to be annually balanced and exhibiting little seasonal variability (figure 3.1). If the inversion is constrained by atmospheric CO₂ alone (mode 1 in figure 3.2), a huge source of 5.1 Pg C yr⁻¹ is diagnosed there, together with an uncertainty of the same magnitude and an unrealistically strong seasonal variability. But already the inclusion of the ocean constraint in mode 2 completely changes the picture: North Africa is now estimated to be annually balanced with a reduction in uncertainty of 42%. The seasonal variability is reduced substantial, and the
3.4. Results and Discussion

resulting, almost aseasonal, flux cycle reflects already clearly the control case. Additional inclusion of land and model constraints does not cause any more changes, except a further reduction in uncertainty for the annual mean. The seasonal pattern is in general agreement with T3L2, although they found a more pronounced Northern Hemisphere type cycle with summer uptake and winter outgassing. Note, however, that their estimation is strongly driven by their use of a monthly land prior, while ours is mainly a result of the strong and twofold (annual and seasonal) ocean constraint, which therefore can replace such a land prior to large parts. The neutral annual flux differs from the source of 0.8 Pg C yr\(^{-1}\) and 1.6 Pg C yr\(^{-1}\) estimated by T3L2 and Ja07, respectively. However, due to the generally large uncertainty in this region all estimates are formally consistent on a 1\(\sigma\) basis.

Tropical Asia is found to be an annual CO\(_2\) sink of -0.8 (\(\pm\)0.7) Pg C yr\(^{-1}\). The annual mean flux is estimated close to zero or weakly positive in modes 1 to 3, before the inclusion of the transport model constraint drives it towards that sink. It is also in mode 4 that the uncertainty is reduced significantly by more than 50%, indicating that in this region the transport error limits annual flux accuracy. The seasonal cycle exhibits a pattern with boreal summer release and winter uptake, however not statistically significant. The annual sink found by the control inversion is significant at an 87% level, compared to the insignificant source and sink estimates of T3L2 and Ja07, respectively.

The land constraint (mode 3) for annual pantropical land flux was -0.1 (\(\pm\)1.1) Pg C yr\(^{-1}\) (Sarmiento et al., 2010). The posterior estimate of 1.1 (\(\pm\)0.9) Pg C yr\(^{-1}\) deviates clearly from that, which is a result of both the atmosphere and ocean constraints (modes 1 and 2 in figure 3.4). Hence, for tropical land areas, and in particular for Tropical America, the information from land inventories and from ocean and atmosphere inversions disagree (figure 3.3). The ocean constraint drives the Tropical American region towards a strong source of carbon, while the tropical regions in Asia and Africa are estimated to be sinks and neutral, respectively. To see whether it is the ocean or rather the atmosphere constraint that implies the tropical source, we conducted an inversion (not shown) where we replaced the twofold ocean constraint by a monthly oceanic CO\(_2\) prior based on pCO\(_2\) measurements alone. This prior had the same seasonal flux pattern as the original one, but with much weaker annual constraining power. As a result, the Tropical American source was halved and the pantropical source decreased to 0.2 Pg C yr\(^{-1}\). Thus, it is clearly the annual constraining strength of the ocean prior, combined with its information about the error structure of regional ocean fluxes, that implies the sizeable tropical land source, and not so much the prior information about the oceanic seasonal cycle.
Figure 3.4 Control fluxes for selected aggregated regions are compared to results obtained in each mode of the joint inversion.
3.4. Results and Discussion

3.4.4 Interrelation of Tropical source and Northern sink

Here we take a look at how the results change if atmospheric models are selected according to Stephens et al. (2007), i.e. how the control results compare to those averaged over all available 120 transport model pairs (mode 3). As outlined in the previous section, those models accurately reproduce Northern Hemisphere vertical CO$_2$ gradients on an annual mean basis. Their success is the result of too strong ventilation both in boreal winter and summer, which tend to compensate for each other. However, such an enhanced ventilation has an effect on seasonal CO$_2$ fluxes both over Northern Hemisphere and Tropical land regions, as shown below.

In an atmospheric model with too strong ventilation in boreal winter, the respired CO$_2$ over or downwind Northern Hemisphere continental areas is removed faster from the boundary layer. In an inversion this requires a stronger carbon source to match the elevated CO$_2$ signal at boundary layer sites during winter. Similarly, in summer a strong ventilation makes a larger carbon sink necessary over Northern land. This correlation between flux strength and ventilation rate is very strong for winter and rather weak for summer (probably due to opposing effects of fossil emissions and terrestrial carbon uptake during summer, cf. Stephens et al. (2007)) and leads to a larger seasonal amplitude of Northern land CO$_2$ fluxes (figure 3.5). If averaged over a year, the net effect of elevated ventilation is to reduce the annual mean carbon sink over Northern Hemisphere land. These two features of enhanced seasonal amplitude and reduced annual sink can be identified in our results for, e.g., Europe and temperate North America (figure 3.2), where all-model mean (mode 3) results are compared with control (mode 4) results. Annual uptake of aggregated Northern Hemisphere land is reduced by more than 0.6 Pg C yr$^{-1}$ (figure 3.4).

As global CO$_2$ mass balance must always be maintained, and because seasonal coupling of atmospheric transport and terrestrial carbon fluxes leads preferentially to the formation of CO$_2$ gradients between northern and tropical land, the reduced Northern Hemisphere land sink is accompanied by a reduced tropical land source. Accordingly, our control inversion estimates a decrease of nearly 0.6 Pg C yr$^{-1}$ in tropical outgassing, when compared to the mode 3 inversion. Figure 3.3 illustrates this interrelation for all 4 inversion modes: the evolution of the tropical source appears mirrored to that of the northern sink.

However, this tropical source reduction is not distributed homogeneously among the continents of Africa, Asia and America. On the one hand, Northern Africa remains neutral and Tropical Asia is estimated to be a sink. Over tropical America, on the other hand, CO$_2$ release increases by almost 0.6 Pg C yr$^{-1}$ to 1.9 ($\pm$0.8) Pg C yr$^{-1}$ between mode 3 and 4. Both for the tropical Asian and American regions the modeled uncertainties are greatly reduced, because the selected models agree better on the flux estimates than all models. The model mean Bayesian error (within-model uncertainty) becomes comparable to the across-model
3.4.5 Implications for the Amazonian carbon balance

To obtain a flux estimate for Amazonia we map the Tropical American source back on the prescribed within-region flux pattern and subsequently re-integrate over the Amazon biome (as delineated after J.C. Riveros-WWF Peru, 6.8 million km$^2$). For better comparison with other studies we follow Malhi et al. (2008) and define Amazonia as the Legal Amazon in Brazil and the Amazon river watershed and Guayanas region outside of Brazil, summing up to 5 million km$^2$. This results in a net annual mean Amazonian source of $1.1 \pm 0.5$ Pg C yr$^{-1}$ during 1992-1996.
Estimates for carbon release in Amazonia due to deforestation lie between 0.3 and 1.1 Pg C yr\(^{-1}\) for the 1990s (Malhi et al., 2008; Ramankutty et al., 2007; Houghton, 2003; DeFries et al., 2002) with a central estimate around 0.5 Pg C yr\(^{-1}\) (Malhi et al., 2008). When using this estimate our results imply a net source of 0.6 (±0.5) Pg C yr\(^{-1}\) in areas of pristine forests. A source of this size disagrees with most bottom-up studies, which usually find pristine tropical forests to be sinks for carbon. Baker et al. (2004) and Phillips et al. (1998) conducted studies based on observed changes in local biomass inventories and suggested that intact forests (plot results upscaled to 5 million km\(^2\)) are a sink of up to -0.6 Pg C yr\(^{-1}\). Upscaled estimates from eddy covariance measurements span the range from -0.5 to -3.0 Pg C yr\(^{-1}\) sink strength in old-growth Amazonian forest (Ometto et al., 2005). Their spread reflects the heterogeneity within tropical forest ecosystems and the associated difficult upscaling process. The study of Saleska et al. (2003) was also based on eddy flux observations, but for the first time identified a carbon source of 0.65 Pg C yr\(^{-1}\) in undisturbed forest areas. As opposed to other flux tower studies they applied a correction to their measured nocturnal fluxes, in order to compensate for the effect of lateral carbon exchange that may not be detected by the instruments during calm nights without turbulence, resulting in an underestimation of nocturnal carbon release. They pointed out that without this correction they would have found a sink instead of a source, as most other flux tower studies. Gurney et al. (2004) estimated in their T3L2 inversion a source of 0.2 Pg C yr\(^{-1}\) for Tropical America (corrected for deforestation).

Thus, the possibility of a positive carbon net flux from Amazonia is not excluded by bottom-up studies. An annual mean 0.6 Pg C yr\(^{-1}\) outgassing such as found by our control inversion is located in the upper part of the range, though consistent with the study of Saleska et al. (2003). We may ask the question if the selection of transport models according to Stephens et al. (2007) is justified, as we would have found a considerably smaller source when averaged over all 120 models (mode 3). Generally, as mentioned above, these models predict less tropical outgassing in conjunction with a weaker Northern Hemisphere land sink. But this reduction in the tropical source is not distributed evenly and is a result of the combination of partly offsetting effects: a significant carbon sink in Tropical Asia and a substantial increase in Tropical American CO\(_2\) release, while northern Africa remains neutral. Due to the large uncertainties the latter is statistically not significant, but both the Tropical Asian sink and the Tropical American source are well constrained, mainly due to a substantial enhancement of model agreement in the subset compared to all model pairs. The posterior error structure does not reveal any significant correlations between the two regions, suggesting that both are constrained individually. Based on a simple analysis of the modeled transport fields, we identified the two stations with dominant influence on the flux estimate of Tropical America. When the data from these stations are excluded from the inversion, we obtain almost no change for the control source. However, the posterior (negative) correlation is now substantial, indicating that, without these observations, the inversion is not able anymore to distinguish between tropical regions in Asia and America. Hence, while the tropical Ameri-
can flux can be separated with the available data, its value is predominantly determined by
the applied constraints. As mentioned above, the driving constraints for the annual Tropical
Asian flux are the model selection and the ocean prior, while for the Tropical American region
it resembles a tug of war between the land constraint and the ocean and model constraints.

To see the influence of the ocean constraint on the Tropical American flux, it is also useful to
look at Ja07’s annual joint inversion, as they constrained the ocean with the same (annual)
flux information. They found an even stronger source of 3.1 (±2.4) Pg C yr\(^{-1}\) across Tropical
America. But due to the large uncertainty they did not make a strong statement for a carbon
releasing Amazon rainforest and focused on pantropical land instead. In their inversion tropi-
cal land fluxes are distributed across continents in a different way: their huge 4.2 Pg C yr\(^{-1}\)
flux results mainly from the allocation of a considerable 1.6 Pg C yr\(^{-1}\) source over northern
Africa, whereas we find this region to be neutral. They also did not see a significant sink over
Tropical Asia.

The seasonal cycle over Tropical America (figure 3.6) suggests the region to be nearly neutral
during months January to July, with an average carbon release of only 0.2 (±1.1) Pg C yr\(^{-1}\).
However, for the remainder of the year (months August to December), a very strong release
of 4.2 (±1.3) Pg C yr\(^{-1}\) is estimated. The August-to-December outgassing is a robust feature
across all modes of our joint inversion (figure 3.2): it is already apparent in the atmosphere-
only inversion, but becomes statistically significant only in mode 4. The transition from a neu-
tral balance to a strong CO\(_2\) source happens rapidly between July and August, and is timed
2-3 months after the onset of the dry season in central and southern Amazonia. The out-
gassing continues for another 2-3 months with only little decline, before the region switches
back to its neutral state in January. As Amazonia represents the main part of the Tropical
American region, we might conclude from these results that large parts of the rainforest re-
lease CO\(_2\) as a consequence of dryer conditions, with a response delay of 2-3 months to the
beginning and end of the dry season. While this is consistent with other studies (e.g. Phillips
et al. (2009)), the uncertainties associated with our monthly flux estimates are too large to
support strong statements about this process. However, despite the large monthly uncer-
tainty, we can state that the mean August-to-December CO\(_2\) source exceeds 2 Pg C yr\(^{-1}\) with
a probability of 95%, and the mean January-to-July flux does not exceed 1 Pg C yr\(^{-1}\) with a
probability of 77%.

3.4.6 Aggregated Tropical and Southern Land

For Southern Hemisphere extra-tropical land the control inversion estimates a sink of -0.6
(±0.6) Pg C yr\(^{-1}\), in good agreement with the T3L2 flux of -0.9 (±1.2) Pg C yr\(^{-1}\). By contrast,
Ja07 estimated a large uptake of -2.4 (±2.0) Pg C yr\(^{-1}\). They assign a significant uptake
also to southern extra-tropical America, which is only a weak sink in both the T3L2 and
3.4. Results and Discussion

Figure 3.6  Black lines show the posterior results from the control (mode 4) inversion for the 11 land regions. According to the sign convention, a positive flux is directed into the atmosphere. Gray shaded areas represent the total posterior uncertainty as one standard deviation, which in turn includes contributions from the Bayesian “within” error as well as from the “between” error arising from transport model spread. Results are compared to the T3L2 (Transcom 3, Level 2) atmosphere-only inversion of Gurney et al. (2004) in green, including their reported total uncertainty.
our studies. It should be noted, however, that estimates for the Southern Hemisphere land flux are generally accompanied by large uncertainty due to the very limited availability of CO₂ observations from Southern Hemisphere continental sites. This makes the various flux estimates formally consistent with each other, despite their large differences.

From figure 3.2 it can be seen that for Southern Africa the ocean constraint plays again a key role in determining the annual mean flux as well as the seasonal variability. An explicit monthly land prior is therefore not needed to constrain fluxes much better than by atmospheric CO₂ observations alone.

The same is true in Southern extra-tropical America, where the ocean constraint massively adjusts the annual mode 1 sink of 4.5 Pg C yr⁻¹ to a nearly balanced flux, which is subsequently only changed marginally by the land and model constraints. The seasonal cycle is also determined to a large extend already by the ocean constraint. In modes 2 to 4 of our joint inversion the annual mean flux is consistent with T3L2. The seasonal flux amplitude is also consistent among modes 2 to 4, but is greatly enhanced compared to T3L2. This is partly due to the monthly land prior used by them, but the main reason is a subtle, but necessary, correction in the compilation of the T3L2 cyclostationary transport response fields (see appendix A for details). While this correction principally affects results for all regions, it is negligible in all but the Southern extra-tropical American region. We reproduced the T3L2 results and found that, if Gurney et al. (2004) had applied this correction in their T3L2 inversion, they would have detected a larger seasonal amplitude as well.

As is often the case in atmospheric inversion studies (Gurney et al., 2004; Jacobson et al., 2007a; Gurney et al., 2002), the posterior error structure exhibits a strong negative correlation between tropical and Southern Hemisphere land areas, allowing the sum of both to be constrained much better than each individually. This follows from the ill-posed character of the inversion problem, resulting from the sparse and very heterogeneously distributed observational network. Global uptake is well constrained by the difference of fossil fuel emissions and the observed global growth rate of atmospheric CO₂. And in our joint setup the oceanic portion of annual global uptake is also well constrained, due to the tight range of estimates from the 10 ocean transport models underlying the ocean inversion. This is a consequence of the vast amount of available ocean interior carbon data fed into the inversion, thereby making the ocean inversion an overdetermined problem with no need for prior information, nor for a global oceanic uptake constraint. The global land flux is then, as a residual, also well-constrained, which leads to a posterior error structure with preferably negative correlations between less well-constrained land regions, such as the Tropical and Southern Hemisphere land areas. As a result, the aggregated flux from Tropical and Southern Hemisphere Land (henceforth, TSL) can be estimated with higher certainty than the individual fluxes (figure 3.7). In figure 3.3 the strong anti-correlation between Northern Hemisphere extra-tropical land and TSL manifests itself in a nearly perfectly mirrored curvature for the two regions.
3.4. Results and Discussion

Our control inversion finds a TSL source of 0.5 (±0.4) Pg C yr$^{-1}$, much smaller than the 1.8 (±1.1) Pg C yr$^{-1}$ estimate of Ja07 and also the 1.0 (±1.3) Pg C yr$^{-1}$ estimate of T3L2. The main reason for this small source is the relatively small outgassing from tropical land, as discussed in the previous sections (figure 3.7). The main portion of the uncertainty reduction is due to the inclusion of the ocean and the model constraints. The TSL seasonal cycle is characterized by strong CO$_2$ release in (boreal) summer and fall, and uptake during winter and spring (figure 3.4). The seasonal amplitude is much larger than estimated by T3L2, which results from the combined effect of increased Southern Land seasonal amplitude and a nearly anti-phased tropical seasonal cycle.

3.4.7 Constraints from atmospheric CO$_2$ column pseudodata

The main goal of our pseudodata inversion was to quantify the potential that real satellite-derived CO$_2$ column measurements have to constrain a synthesis inversion for the estimation of sinks and sources of CO$_2$ at the Earth’s surface, in particular when compared to atmospheric surface CO$_2$ measurements and carbon measurements from the ocean.

Results are shown in figure 3.8. Before we turn into discussing the column data constraint, there is already an interesting result from the pseudo-inversion based on the surface network data: if the ocean constraint is included, the median of the posterior flux uncertainty is reduced by about 50% (red lines in the figure), reflecting the strength of the ocean prior. Similar reductions in uncertainty were at the upper edge of what we found in our previous real data inversions. This is because the transport error vanishes in the pseudo-inversion, while in the real data inversion it is an important source of error that cannot be reduced as much as the
internal (within-model) error by the ocean constraint. Note that the posterior uncertainties are generally small here, because they represent only the within-model error contribution.

**Figure 3.8** Annual posterior flux uncertainty for a given column data uncertainty for the pseudo-inversions. Red curves represent the pseudo-inversion constrained by PBL data at existing station locations. Blue curves represent the inversion of the column pseudo data sampled at each pixel of the underlying transport model. Black curves represent the inversion of the column pseudo data sampled at only the station locations. Dashed lines show the results if only the pseudodata are used to constrain the inversion, while solid lines show the results if the ocean prior is used in addition. Yellow markers indicate important features described in the text.

Blue curves in the figure represent the posterior flux uncertainties from the pixel-based column inversion. As the prescribed column data uncertainty decreases, the posterior flux estimates become better constrained. The interesting point here is the crossing point of the blue and red curves, because that point gives the column data uncertainty needed to meet the constraining power of the surface data. For the non-joint inversion without ocean prior, this threshold is around 2 ppm. Such a threshold is in very good agreement with estimates from previous studies of 1.7 to 2.5 ppm (e.g. Rayner and O’Brian (2001); Kadygrov et al. (2009)) and lies between 0.5% and 1% accuracy of the column data. Accuracies in this range are feasible for some of the active or proposed satellite projects (see table 2.2), whereas accuracies above 0.5% would be difficult to achieve with passive instruments.

If the ocean constraint is taken into account, the accuracy threshold is reduced to about 1.2 ppm, or about 0.3 to 0.4%. Such accuracy may only be achieved by active sensing (Breon and Ciais, 2010), yet no active sensing mission is scheduled to launch before 2016. If the above analysis is based on monthly posterior flux uncertainty (instead of annual), the inclusion of the ocean constraint has only a minor effect (not shown here). This is an anticipated finding, because the ocean constraint is strong on an annual basis only (section 2.3) due to the longterm character of the ocean inversion.

The black curves give the results for a column data inversion that does not exhibit global
3.5 Conclusions

Results of a joint inversion are presented in this paper, estimating 1992-1996 monthly mean CO₂ fluxes for 22 land and ocean regions. The inversion is performed in 4 modes, in each mode including an additional constraint. These constraints are 1) atmospheric CO₂ observations, 2) ocean DIC and pCO₂ data, 3) land priors for selected regions, and 4) transport model selection. The setup generally follows a Bayesian synthesis inversion scheme with the particular feature that in mode 2 not only sea-to-air flux priors are considered, but also the full error structure (variance-covariance information) as obtained from an ocean interior inversion of DIC. This makes our approach equivalent to using a one-step Kalman filter approach like Jacobson et al. (2007a). No further constraints are considered, in particular no monthly land prior information for all regions.

The inclusion of DIC- and pCO₂-based information allows us to estimate a set of sea-to-air fluxes that is concurrently consistent with these datasets. Major differences compared to atmosphere-only inversions, such as the one of Gurney et al. (2004) or our mode 1 inversion, include the finding of an outgassing across all tropical ocean basins, as well as an uptake across ocean basins in the Southern Hemisphere. As a result from the data-based prior, seasonal variations are small in these regions, preventing the inversion from misallocating terrestrial seasonal signals to oceanic regions, which often occurs in atmosphere-only inversions. From the set of well-constrained sea-to-air fluxes information is propagated on land, making joint estimates for land-to-air fluxes consistent with the oceanic datasets, too.

The ocean data constraint is strong, particularly on an annual flux basis. Together, the constraints of atmospheric CO₂ and the oceanic carbon data require other data sources to have a high information content (with regard to surface flux estimation) if supposed to compete with them. For example, our experimental inversion of simulated atmospheric CO₂ column concentrations suggests that satellite-derived column mean concentrations need to be retrieved with a minimum accuracy of 1.2 ppm in order to provide the same information on annual CO₂ fluxes as the combined data streams from the ocean and lower atmosphere.
The required accuracy reduces to 2 ppm if only the lower atmospheric data from the surface CO$_2$ network are considered.

Generally, the power of each constraint in the joint inversion is substantial: When one constraint is omitted, the change in estimated fluxes and uncertainties is not negligible (figure 3.3). For some regions, such as Northern Hemisphere land, all constraints drive the results in the same direction. In other regions, two or more constraints point into opposite directions, for example in Tropical America, where the ocean prior and selected models imply a strong source, while the inventory-based annual land prior suggests a neutral balance. Uncertainty is reduced significantly by each constraint, though the sign significance of flux estimates does not necessarily increase, due to the regularizationary effect of the constraints (figure 3.3 and table 3.2).

Inclusion of the ocean constraint has a major impact on flux estimates from under-constrained regions with only few available CO$_2$ measurements, located in the Tropics and Southern Hemisphere. In particular, it is the key constraint for the regions in Northern and Southern Africa and South America. Here the inclusion of further constraints does not lead to major changes in the seasonal flux cycle. It also does not significantly alter the annual mean flux, but further decreases uncertainty.

The seasonal cycle in temperate Asia is mainly shaped by the inclusion of the ocean constraint as well, despite the better coverage of atmospheric CO$_2$ sites there. We find the growing season to last only from May to July, in contrast to estimates from our atmosphere-only inversion and the T3L2 results of Gurney et al. (2004), which extend it until September. The region loses its status as an annual mean carbon sink and is now estimated to be neutral or a weak source.

Fluxes from Northern Hemisphere land are already well-constrained by atmospheric CO$_2$ alone, so that our results are in good agreement with previous atmosphere-only inversions. We estimate, however an enhanced seasonal amplitude, driven by stronger winter respiration. This is mainly a result of the transport model constraint in mode 4, because selected models are characterized by enhanced ventilation over Northern Hemisphere land. Due to the seasonal coupling of atmospheric transport and CO$_2$ fluxes, those models require stronger CO$_2$ fluxes, in particular during winter when this coupling is strongest.

Pantropical land is estimated to be a source of carbon, significant at the 90% level. This source is located solely in Tropical America. It is partially offset by a sink in Tropical Asia. The transport model selection has a strong influence on tropical fluxes, as it connects the Tropics with the Northern Hemisphere through the models’ ventilation schemes. As a result, the model constraint in mode 4 not only causes reduced Northern land uptake, but also a reduction of Tropical land source. While for pantropical land this reduction is obtained, the model constraint influences tropical fluxes from the American and African/Asian continents.
3.5. Conclusions

differently. Indeed, the aggregated African and Asian tropical flux is decreased, in fact turns from a source into a sink. However, the Tropical American source increases, despite the fact that the sink over Temperate North America is reduced. This suggests that over the American continent the Northern-Tropical coupling does not play a key role, as opposed to the rest of the globe. Instead, it seems that those models show a stronger coupling between the Tropics and the Southern Hemisphere, as the South American flux is decreased (i.e. turns from a source into a sink) by the model constraint.

For the Amazonian region we find an annual mean biospheric source of 0.6 Pg C yr\(^{-1}\) after downscaling and subtracting a typical deforestation rate. This appears inconsistent with most bottom-up studies, which typically assign tropical American rainforests a sink of carbon. However, Saleska et al. (2003) pointed out that analyses based on eddy covariance towers could be biased without proper correction of nocturnal fluxes during calm nights. After them applying such a correction, they obtained a source of 0.65 Pg C yr\(^{-1}\) for the Amazonian rainforest, in very good agreement with our result.

We estimate a small source for aggregated Tropical and Southern Hemisphere Land (TSL) of 0.5 (±0.4) Pg C yr\(^{-1}\), as a result of near cancellation of a pantropical source and Southern Hemisphere sink. The TSL region is strongly anti-correlated with Northern Hemisphere extra-tropical land, because of the well-constrained global land sink. This allows the TSL flux to be estimated with such small uncertainty. Compared to the T3L2 study of Gurney et al. (2004) we find a much higher seasonal amplitude over TSL, resulting from an anti-phased tropical cycle combined with increased seasonal amplitude over Southern Hemisphere extra-tropical land.

The more pronounced seasonal cycle over TSL, in concert with the increased winter respiration over Northern Land, leads to significant changes in the seasonal cycle of the global land area (figure 3.4). Global seasonal amplitude is increased and the seasonal pattern markedly asymmetric compared to the T3L2 results. The asymmetry results from large release rates in September-December, together with small uptake rates in January-April. Such an asymmetry is in agreement with results from the ORCHIDEE vegetation model (CARBOSCOPE).
Chapter 4

Decadal trends of regional carbon fluxes from a joint ocean-atmosphere inversion

4.1 Abstract

We show decadal trends in seasonal non-fossil land-atmosphere and ocean-atmosphere carbon exchange from a joint ocean-atmosphere inversion for the three periods 1980-89, 1990-99 and 2000-08. In addition to atmospheric CO$_2$, constraints from ocean interior DIC and surface ocean pCO$_2$ are included, yet no explicit land prior. The sensitivity of results to observational network composition and selected atmospheric transport models is explored. We find the terrestrial biosphere to contribute 87% (-0.048 Pg C yr$^{-2}$) to the trend in the overall removal of carbon from the atmosphere, whereas the oceans’ share is only 13% (-0.007 Pg C yr$^{-2}$). Land uptake exhibits a distinct intensification in the 1990s compared to the 1980s of 0.75 ($\pm$0.37) Pg C yr$^{-1}$, driven primarily by southern hemisphere land, particularly the southern half of the African continent. The northern extratropics act as increasing carbon sink, though their relative contribution to global land uptake reduced by about 20% over the examined period. Boreal regions account for 79% of the northern extratropical sink trend, induced by enhanced growing season net uptake in boreal America and declining dormant season net release in boreal Asia. Tropical land is estimated to act as an increasing source of carbon, with source magnitude and trend being dominated by intensified outgassing in tropical America during the Amazon-mean wet season. Trend estimates for tropical America are always positive, irrespective of network choices and the number of applied constraints. However, between the 1990s and 2000s the tropical American source weakened by 0.2 Pg C yr$^{-1}$, in broad agreement with an observed decline in deforestation-induced fire emissions in the Amazonian region. By contrast, aggregated regions in tropical Africa and Asia act as carbon sink with little discernible trend. Among all constraints the ocean prior plays a key role in setting the trends within latitudinally aggregated bands. At the regional level it has a major
impact on carbon exchange for all tropical regions and southern Africa, but also for observationally better constrained regions in North America and temperate Asia. The European trend exhibits an unexpected sensitivity to network composition.

4.2 Introduction

Over the recent decades the fossil fuel CO$_2$ emissions to the Earth’s atmosphere have been steadily increasing from 5.5 Pg C yr$^{-1}$ in the 1980s and 6.4 Pg C yr$^{-1}$ in the 1990s to 7.7 Pg C yr$^{-1}$ during 2000-08 (Boden et al. (2010); cf. figure 4.1). At the same time, emissions from land-use change remained relatively constant around 1.5 Pg C yr$^{-1}$ (Le Quere et al., 2009). The oceans and terrestrial biosphere removed slightly more than half of this extra carbon from the atmosphere, whereas the rest led to the observed increase in atmospheric CO$_2$ concentration. The quantification as well as decadal trends in the ocean and land sinks are crucial for projections of their future behavior, which in turn allows for projections of future atmospheric CO$_2$ concentrations and associated climate change (Sitch et al., 2008; Le Quere et al., 2009).

Carbon exchange between the land and atmosphere is of particular interest, because it is more directly influenced by anthropogenic activity than the ocean and is hypothesized to be prone to large-scale carbon-climate feedbacks (e.g. Potter et al. (2003)). Although land uptake exhibits large interannual variability, e.g. in response to ENSO and volcanic eruptions, the identification of a decadal trend would support our understanding of the terrestrial carbon cycle and potential future changes. Direct approaches to estimate the terrestrial sink include eddy covariance measurements (e.g. Hutyra et al. (2007)) and inventory studies (Phillips et al., 2009; Lewis et al., 2009; Ciais et al., 2010). Both methods are extremely valuable to accurately estimate local to regional carbon exchange, but they represent only local vegetation type and are, hence, difficult to upscale to continental or global regions characterized by heterogeneous biome structure. An alternative are biosphere models, which can be used to simulate present day as well as trends in the land carbon sink (Potter et al., 2003; Le Quere et al., 2009; Piao et al., 2009). However, the underlying drivers - such as climate, CO$_2$ and land-use change - are afflicted with uncertainties that lead to a considerable spread of results. Other methods include the indirect observation of carbon-related measures from space, such as LAI or EVI (Myneni et al., 2007; Huete et al., 2006). It remains a challenge, however, to translate those indices into net carbon exchange.

Alternatively, one can use the results of atmospheric CO$_2$ inversions to estimate fluxes and trends at regional and global scales. The basic principle of each inversion is to combine information from atmospheric CO$_2$ measurements with simulated atmospheric transport to obtain carbon fluxes at various spatiotemporal resolutions (Enting, 2002). Such inversions
have been shown to be well suited to estimate fluxes at continental scales with monthly resolution, limited mostly by model resolution as well as measurement coverage and frequency (Gurney et al., 2004, 2008; Baker et al., 2006). In addition to atmospheric CO$_2$ they are constrained by combined ocean and land carbon uptake derived by difference of fossil emissions and global atmospheric CO$_2$ growth rate, which are both known with high confidence (figure 4.1). The partitioning into regional fluxes from ocean and land is then done during the inversion process, usually supported by utilizing *a priori* information about annual mean flux and seasonal cycle for land regions Gurney et al. (2004); Gurney and Eckels (2011). Such land priors have, however, a major influence in regions poorly constrained by observations, such as the tropics. They can, hence, be expected to dominate also any estimated trends from those regions and to obscure the weaker atmospheric CO$_2$ constraint. While they play a dominant role, bottom-up prior information about trends in land uptake at continental level are hard to find in the literature, except for a few regions like Europe or North America. Gurney and Eckels (2011) used model-based land priors based on the Transcom series of inversions to overcome this issue. In order to avoid the uncertainties associated with biosphere models and to rely more on observational constraints, we decided to abstain from land priors and to put more emphasis on observation-based oceanic flux constraints, which then indirectly constrain land fluxes through atmospheric transport.

Our approach is related to the joint atmosphere-ocean inversions of Jacobson et al. (2007a) and the one described in chapter 3, but extends the results to the estimation of decadal trends over the recent three decades, i.e. the 1980s, 1990s and 2000-08. The applied ocean prior is derived from an inversion of ocean interior DIC in combination with surface ocean pCO$_2$ measurements. It provides not only regional ocean-atmosphere carbon fluxes, but also associated uncertainty and spatial correlation structure. When included in the joint inversion, all these information are propagated to the land, making land flux estimates consistent with oceanic carbon measurements and our knowledge about ocean circulation. Interannual variability in oceanic carbon exchange is restricted to their response to atmospheric CO$_2$ accumulation without considering changes in circulation. While this introduces errors, the interannual variability is significantly smaller over the ocean than over land (Bousquet et al., 2000; Le Quere et al., 2009; Sarmiento et al., 2010). Hence, the advantage of constraining the ocean rather than directly the land is the more robust knowledge we have about oceanic carbon trends.

In this paper we study the effect the inclusion of the ocean prior has on trends of land carbon sinks and sources. We also compare our results to the recent interannual atmospheric CO$_2$ inversion of Gurney and Eckels (2011), which include land priors similar to those used in the Transcom inversions (Gurney et al., 2002, 2004; Baker et al., 2006) and based on a land ecosystem model (CASA). We show that such land priors can lead to the estimation of unrealistic ocean trends, for example a weakening of global ocean sink strength that is unlikely to have happened.
Figure 4.1  Annual CO$_2$ uptake by global oceans and land. The uptake is calculated by difference from fossil fuel emissions and observed global growth rate in atmospheric CO$_2$. The latter was derived from time series measurements at Mauna Loa and South Pole stations. The curve represents the station mean and the shade their spread. The mean global uptake for the three inversion periods is given as dotted lines. Note that these values are based on a fit over each period as a whole (cf. methods section), and may not equal the mean of the annual values exactly. The ocean uptake represents the weighted mean result from the ocean interior inversion (see Mikaloff Fletcher et al. (2006, 2007) for details on model weighting). It is corrected for outgassing due to carbon input by rivers. The net land-atmosphere flux is the difference between the global uptake and the ocean sink.
We include three constraints in our inversion by sequence: mode 1 includes only atmospheric 
$\text{CO}_2$, mode 2 additionally includes the ocean prior and mode 3 selects three atmospheric 
transport models that show most skill in simulating vertical $\text{CO}_2$ gradients (Stephens et al., 
2007). For each mode and decade, inversions are performed on two different atmospheric 
observational networks: a common network containing the same stations throughout 1980-
2008 and an individual network specifically designed for each decade. Results represent 
decadal mean monthly carbon fluxes between the atmosphere and 22 oceanic and land 
regions (according to the Transcom regions definitions). The control inversion is defined on 
the common network and includes all constraints.

4.3 Methods

The joint inversion sequentially interprets observations in the ocean and atmosphere to esti-
mate carbon fluxes for 22 regions globally and for three decades 1980-1989, 1990-1999, and 
2000-2008. Sets of 10 oceanic and 11 atmospheric transport models are used to link surface 
fluxes to observations in the respective compartment. The approach is very similar to that 
used in chapter 3, except for differences in observational networks as well as adjustments 
concerning the different inversion periods. Details are explained in the following subsections.

For the oceanic part, ocean interior DIC (Dissolved Inorganic Carbon) measurements from 
a global dataset are inverted to yield longterm mean sea-air fluxes. These estimates are 
augmented with seasonal flux information derived from differences in p$\text{CO}_2$ (the partial pres-
sure of $\text{CO}_2$) in the surface ocean and the atmospheric layer above (Takahashi et al., 2009b), 
in order to form monthly flux estimates averaged over the 1980s, 1990s or 2000s period. 
Uncertainties and error covariances are obtained along with the fluxes.

Monthly mean atmospheric $\text{CO}_2$ observations from the Globalview-co2 (2009) dataset are 
used for the atmospheric inversion. Data are averaged over the three decades and inverted 
to estimate monthly and decadal mean fluxes. Two different kinds of network are designed: 
a common network with 97 sites for all periods, and free networks that are individually set up 
for each period. This was done to assess the robustness of estimated decadal flux trends 
against changes in network composition.

Results are presented for three inversion modes:

1. atmosphere-only inversion,

2. joint ocean-atmosphere inversion, where the sea-air flux and error information from the 
ocean inversion enters the atmospheric inversion as prior constraint, and
3. joint ocean-atmosphere inversion with a restricted set of atmospheric transport models (Stephens et al., 2007).

Each mode is performed on both network types.

### 4.3.1 Observational networks for atmospheric CO$_2$

The selection of atmospheric CO$_2$ stations from the NOAA Globalview dataset is based on criteria related to the ratio of real measurements made at the stations to smoothed and interpolated data. This ratio is evaluated on a monthly timescale as well as over the whole time period of inversion, as described in section 2.2.2. The minimum requirements used in this study equal those used in chapter 3, except that we reduced the minimum required total ratio of real data to all data from 70% to 50%. This is to allow for a few more stations, as otherwise the network for the 1980-89 decade would become too small to ensure robust inverse results (issue of over-fitting, as described below).

![Locations of atmospheric CO$_2$ stations](image)

**Figure 4.2** Locations of atmospheric CO$_2$ stations for all networks considered in this study. Thick gray lines in the background delineate the 22 Transcom regions, consisting of 11 oceanic and 11 land regions.

For each inversion period an individual network was compiled according to these selection criteria ("free" networks in figure 4.2). The sizes of the networks range from 32 sites for the
1980-89 period to 122 sites for the 2000-2008 period, owing to the increasing observational coverage for atmospheric CO$_2$ since the 1980s.

From previous studies it is known that a network size of 70-80 is sufficient to provide robust flux estimates for the 22 Transcom regions (4.2), if the flux resolution is annual or monthly: see annual Transcom 3 inversion of Gurney et al. (2003), monthly cyclostationary Transcom 3 inversion of Gurney et al. (2004), monthly interannual Transcom 3 inversion of Baker et al. (2006), or monthly cyclostationary inversion in chapter 3. For instance, the monthly cyclostationary inversion for 22 regions described in chapter 3 solves for $22 \times 12 = 264$ variables, and is constrained by monthly mean data from 85 sites, thus by $85 \times 12 = 1020$ equations, plus any extra constraint that may be included (such as a global growth constraint). Hence, the problem is formally over-determined (though ill-posed, see Tarantola (2005b); Enting (2002)) and its solution stable. Our free network for the 1980-89 period contains only 32 stations. The number of equations ($32 \times 12 = 384$) approaches the number of variables, making the inversion susceptible to over-fitting issues, as shown in the left panel of figure 4.3. Very small data mismatches are produced compared to the assigned prior data uncertainty. This makes posterior flux estimates very sensitive to small changes in the data and data uncertainty, including changes due to the applied weighting scheme.

**Figure 4.3** The normalized mismatch between posterior and prior data is plotted against prior data uncertainty (1 standard deviation). The left panel shows the unconstrained inversion (mode 1), while in the right panel the ocean constraint is included (mode 2). For each data point $i$, where $i$ represents a combination of station and month, the absolute mismatch $(Mx)_i - d_i$ was divided by the prior data uncertainty $\sigma_{d,i}$ to yield the normalized mismatch $((Mx)_i - d_i)/\sigma_{d,i}$. $M$ indicates the transport matrix, and the vectors $x$ and $d$ contain the flux variables and data points, respectively. The closer a value is to zero, the smaller the mismatch is relative to the prior uncertainty and the better the inversion was able to reproduce the station data accurately. Values between -1 and 1 (within the area framed by the dotted lines) indicate a mismatch that is within one standard deviation from the prior data.
The instability issue occurs only for the 1980-89 inversion with 32 sites. The additional inclusion of the ocean constraint, which is described in the following section, resolves the issue (right panel in figure 4.3).

In addition to the individual networks we also designed a common network for all inversion periods. The main advantage of a common network is that trends in the flux estimates cannot be influenced by changes in network composition, but are only caused by trends in the data themselves. On the other hand, individual networks are preferable if the focus is on absolute fluxes for a single inversion period, as each individual network is specifically designed for that period. In this paper both network types are used in parallel to assess the robustness of absolute fluxes as well as of decadal trends.

The most straightforward approach for a common network for all three periods was to use the intersection of stations of the individual networks. But this reduced the number of sites to 25 (figure 4.2) and made the over-fitting issue - which existed already for the 32 site individual network - more serious (red symbols in figure 4.3). We therefore created another common network, based only on the intersection of the 1990-99 and 2000-08 individual networks. This results in a 97 site network as depicted in figure 4.2, which is used as common network throughout this paper. In order to use this network also in the 1980-89 inversion we increased data uncertainty for sites not included in the 25 site network. For each site this adjustment is based on the site’s interannual variability, see section 4.3.2.

### 4.3.2 Data handling and atmosphere inversion

In the atmospheric part of the joint inversion, a set of 11 atmospheric transport models is used to translate monthly carbon fluxes from 22 regions into global CO$_2$ concentration fields. The regions represent those introduced by the Transcom 3 inversion intercomparison (e.g. Gurney et al. (2003, 2004); Baker et al. (2006)) and consist of 11 oceanic and 11 terrestrial regions (figure 4.2). From all transport models that participated in the Transcom project we chose those for which all necessary response fields were available (see table 4.1 for more information on model characteristics).

For each region and month a pulsed dye tracer flux with prescribed strength was released and propagated through each transport model to obtain the atmospheric concentration response. The modeled response fields were then sampled at the station locations in the applied network. The model-data mismatch is minimized using least squares (section 2.2). As in chapter 3 we report posterior flux uncertainties representing the combined within-model and between-model components.

Observations at the stations are derived from the Globalview dataset. Data processing and
uncertainty assignment for each inversion period are described in section 2.2 and is equivalent to the procedure used in chapter 3 for the 1992-96 period: decadal mean, monthly data are compiled from the Globalview data, which get assigned uncertainty estimates based mainly on the monthly and total ratios of real measurements to smoothed and interpolated data (section 2.2.2). Resulting observational uncertainties are adjusted by a global constant such that an overall model mean reduced $\chi^2$ value close to 1 is achieved.

We applied exceptions to the data processing for the 1980-89 inversion on the common (97 sites) network, in order to account for the fact that some of the sites did not fulfill the ratio criteria, but were nevertheless included for two reasons: 1) to avoid the mentioned issue of overfitting, and 2) to reflect the fact that we actually have some knowledge about what concentrations would have been measured at these stations, based on data values derived from the latitudinal reference boundary layer concentrations. Of course, these reference data vary only latitudinal and are not site-specific. However, by extrapolation of site-specific interannual variability for time periods when measurements were taken, into the 1980-89 period, we have a basis for an uncertainty estimate. The uncertainty is estimated for each month based on the total range of interannual variability over the whole Globalview time period. This results in rather large and therefore conservative uncertainty estimates for these sites with typical values of 5 to 20 ppmv. Data from these stations therefore provide only weak constraints on the inversion, while ensuring its stability. We think that using weak observational constraints from non-direct measurements is preferable over an unstable inversion.

The special setup for the 1980-89 inversion on the common network leads to very different data uncertainties for sites that fulfill the ratio criteria (mostly between 0.5 and 1 ppmv, cf. figure 4.3) and those that do not (5 to >20 ppmv). Because of this, the model mean reduced $\chi^2$ value of the inversion becomes much smaller than 1 for a typical adjustment constant $A$. The usual iterative scaling procedure would decrease data uncertainty (by increasing $A$) until $\chi^2 = 1$ is achieved. This scaling would affect mainly the sites with large uncertainties, and hence dilute the intended distinction between them and sites with real data. To avoid that we did not apply the iterative procedure in this case, but used the same value for $A$ as for the free network inversion. Resulting values for $\chi^2$ are between 0.3 and 0.4, depending on the inversion mode, i.e. unconstrained or with ocean or model constraint. An example for a northern hemisphere data station is shown in figure 4.4 (Globalview station code is zep_01D0).

The effects of CO$_2$ emissions from fossil fuel burning, cement manufacture and gas flaring are removed from the data before the inversion by propagating them through the transport models and subtracting the response in atmospheric CO$_2$ (section 2.2.3). Annual emission estimates are based on Marland et al. (2008) for each year in the inversion period. Averaged emissions over the inversion periods are 5.5 Pg C yr$^{-1}$, 6.4 Pg C yr$^{-1}$, and 7.7 Pg C yr$^{-1}$ for 1980-89, 1990-99, and 2000-08, respectively.
Chapter 4. Decadal trends in carbon exchange

The only additional constraint in our atmosphere inversion (mode 1) is the global carbon uptake by ocean and land. It is derived as the difference between fossil fuel emissions and the global atmospheric growth rate. The growth rate along with uncertainty is computed as described in section 2.2.4. Results imply a global uptake of $2.26 \pm 0.12$ Pg C yr$^{-1}$, $3.10 \pm 0.04$ Pg C yr$^{-1}$, and $3.37 \pm 0.10$ Pg C yr$^{-1}$ for the 1980-89, 1990-99, and 2000-08 periods (cf. figure 4.1).

Posterior flux estimates from the atmosphere-only inversion represent monthly, period-mean carbon fluxes from the 11 oceanic and 11 terrestrial Transcom regions. They are constrained by monthly atmospheric CO$_2$ data and by an annual global uptake estimate. No additional constraints are applied in this mode, in particular no ocean or land prior fluxes. Results are
4.3. Methods

reported as net fluxes that exclude fossil fuel emissions.

4.3.3 Ocean constraint

For mode 2 of the joint inversion, carbon fluxes are not only constrained by atmospheric CO$_2$ measurements, but also by observations of DIC (Dissolved Inorganic Carbon) in the interior of the global oceans. DIC observations are interpreted by an ocean inversion to estimate carbon fluxes across the air-sea interface. Fluxes, as well as the full error and covariance information are included as Bayesian priors in the atmosphere inversion to yield jointly constrained flux estimates for the 22 land and ocean regions. These estimates are consistent with three independent and complementary carbon datasets: atmospheric CO$_2$, ocean interior DIC and surface ocean pCO$_2$ (see sections 2.3 and 2.4).

Results from the ocean inversion are strong constraints for annual sea-air fluxes averaged over a decadal to multidecadal timescale, however they do not provide seasonal flux variability. In this study, the ocean inverse fluxes are augmented with information about seasonal flux patterns based on surface ocean pCO$_2$ measurements (Takahashi et al., 2009a). The result is a composite prior for the joint inversion that exhibits strong constraining power on the annual basis including uncertainties and correlation structures. For seasonal flux patterns, its strength is significantly smaller. Typical flux uncertainties are 0.06 and 0.5 Pg C yr$^{-1}$ for the annual and monthly timescales, respectively.

The computation of the monthly, cyclostationary ocean prior generally follows the explanations in sections 2.3 and 2.4, but the inverse anthropogenic air-sea flux estimate is adapted to account for the different inversion periods as described below. The ocean inversion is a classical least squares synthesis inversion, where a set of modeled footprints is superposed linearly until the resulting tracer field most closely matches observations. The inversion assumes the ocean transport model to be perfect, i.e. any posterior flux uncertainty results only from observation errors and limited observational coverage. A model error cannot be quantified during the inversion. To estimate the model error we follow the same approach as for the atmosphere inversion: a suite of 10 transport models with different mixing schemes is applied, and the spread in results among them is used as between-model error contribution. In conjunction with the 11 atmospheric transport models we get a total number of 110 model pairs underlying the joint inversion mode 2.

The ocean inversion separately estimates preindustrial (Mikaloff Fletcher et al., 2007) and anthropogenic (Mikaloff Fletcher et al., 2006) fluxes. Observations to be interpreted are derived from the global GLODAP database (Key et al., 2004), enriched with data from selected historical cruises. The anthropogenic tracer is computed by the $\Delta C^*$ method of Gruber et al. (1996), and the preindustrial tracer results from the difference between observed DIC and
the anthropogenic carbon. Both tracers are inverted using the same transport models and methodology. For the purpose of this study they are eventually combined to form an estimate of net sea-air flux, which serves as prior in the joint inversion.

Because ocean inverse results represent longterm mean fluxes, they cannot easily be referenced to individual years. However, a first order approach to accomplish that was developed by Gloor et al. (2003) and relies on the fact that the preindustrial part of the net flux by definition does not change over time, and that the change of the anthropogenic flux is strongly related to the evolution of atmospheric CO$_2$ concentration. Under the assumptions of a stationary ocean circulation and near-exponential growth in atmospheric CO$_2$, Mikaloff Fletcher et al. (2006) and Gloor et al. (2003) show that a temporal scaling $\phi$ can be applied to reference the longterm mean anthropogenic flux to individual years,

$$\phi(t) = \frac{x_{CO_2}(t) - x_{CO_2}^{PI}}{\int (x_{CO_2}(t) - x_{CO_2}^{PI}) dt},$$

(4.1)

where $x_{CO_2}(t)$ is the atmospheric CO$_2$ mixing ratio at time $t$ and $x_{CO_2}^{PI}$ indicates the preindustrial value, assumed to be 280 ppm. $\phi$ is evaluated for each year in the desired joint inversion period, and the scaled anthropogenic flux is added to the preindustrial flux to yield an annually referenced total net flux.

When combining an ocean with an atmosphere inversion, the riverine carbon loop needs to be handled consistently, i.e. the fraction of carbon that is taken up over land, then reaches a river and is transported into the ocean, where it is eventually released to the atmosphere. That part of the riverine carbon reaching the open ocean (coastal regions are not resolved by the ocean models) will be misinterpreted by the ocean inversion as uptake across the air-sea interface. We apply a regional correction to the posterior ocean inverse flux according to section 2.3.2; the correction sums up globally to 0.45 Pg C yr$^{-1}$. In figure 4.1 the resulting annually referenced and riverine corrected global ocean flux is shown.

Information about the seasonal flux cycle is based on the pCO$_2$ dataset of Takahashi et al. (2009a). In their study, Takahashi et al. (2009b) describe how they computed a monthly sea-air carbon flux climatology on a 1x1 degree grid, based on the pCO$_2$ difference in the surface ocean and the adjacent atmospheric layer. They applied a globally constant trend of 1.5 $\mu$atm yr$^{-1}$ to surface ocean pCO$_2$ to reference their climatology to the year 2000. To translate pCO$_2$ differences into fluxes they used a specific gas exchange model, which is based on a quadratic windspeed parameterization with parameters that lie within the range found in the literature. For more details on the exact formulation of the gas exchange model, see Takahashi et al. (2009b) or section 2.4, where a tabular overview of existing models is also provided (table 2.1).

We integrated the pCO$_2$-based flux estimates over the 11 oceanic regions and adjusted
them such, that their annual mean matches our ocean inverse estimates. Flux uncertainties were computed based mainly on 1) the \( \text{pCO}_2 \) measurement density in each region, and 2) the spread in flux results when applying seven additional gas exchange models. For further details on the uncertainty assignment, we refer to section 2.4.

General characteristics of the resulting composite ocean flux prior are small annual and sizeable monthly uncertainties. Annual fluxes from the ocean inversion are preferred over annual \( \text{pCO}_2 \)-based estimates, as they are consistent with a large amount of ocean interior DIC data and do not depend on the formulation of gas exchange at the ocean's surface. In addition, the ocean inversion provides errors as well as correlations. On the other hand, ocean inverse fluxes are based on the assumption of an invariant ocean circulation, and their interannual variability is computed according to a simple model based on atmospheric \( \text{CO}_2 \) evolution only. The \( \text{pCO}_2 \)-based seasonal flux cycle is referenced to year 2000. No attempt was made to compile an interannually varying seasonal cycle, because the temporal coverage of the \( \text{pCO}_2 \) data is not sufficient to provide a robust basis.

### 4.3.4 Joint Inversion

The joint inversion combines information from the composite ocean fluxes with atmospheric \( \text{CO}_2 \) measurements in a Bayesian framework (see section 2.5). The setup consists of three different modes, each incorporating one more constraint: 1) atmospheric \( \text{CO}_2 \) concentrations, 2) ocean interior DIC and surface ocean \( \text{pCO}_2 \) data, and 3) atmospheric transport model restrictions.

As outlined before, monthly ocean prior uncertainties are moderate, compared to the small annual uncertainties from the ocean inversion. The latter are included as explicit constraints, i.e. as additional "observations" as described in section 2.5. The atmospheric global growth constraint is included in the same manner (section 2.2.4).

Along with the posterior fluxes, a joint posterior covariance matrix is computed, representing the within-model covariance. Joint estimates are computed for each of the 11 atmospheric transport models listed in table 4.1 using ocean prior flux and error estimates from each of the 10 ocean inversions (ocean transport models also listed in table 4.1). This gives 110 joint flux vectors and within-model covariance matrices for each network and inversion period. The spread among this ensemble is added to the within-model covariance as squared sum, resulting in a total covariance matrix. Flux uncertainties reported in this paper always refer to that total estimate.

The model mean reduced \( \chi^2 \) value is generally close to 1, due to the iterative procedure described in section 4.3.2. An exception was made for the 1980-89 inversion on the common
(97 sites) network, in order to maintain the desired heterogeneous uncertainty distribution among stations fulfilling the ratio criteria and the remaining ones.

For our control inversion (mode 3) we reduced the number of atmospheric transport models to three, which leaves 30 model pairs underlying the inversion’s statistics. The selection of models follows the recommendation of Stephens et al. (2007). It is based on their assessment of the models’ ability to reproduce observed vertical gradients of CO$_2$ in the atmosphere (see section 2.5.2). In a recent study, Pickett-Heaps et al. (2011) highlight one of the three selected models, i.e. the TM3 model. It performs better by some measures than other models to reproduce vertical gradients at two stations: Cape Grim, Tasmania, and Carr, Colorado.

The following results and discussion section is centered on a set of "control" results for each inversion period. These represent the joint inversion, mode three, on the common network. That is, control estimates include the composite ocean constraint and are restricted to the Stephens’ subset of models. However, influences of the ocean and model constraints, as well as the various networks (figure 4.2), are also explored.

### 4.4 Results and Discussion

In this section we present the results obtained from our control inversion (mode 3 on the common network) and discuss them in the context of previous studies. Additionally we investigate the sensitivity of the control results to the choice of network as well as to the inversion mode. Two types of network are considered: those individually designed for each period (individual networks) and the common network with 97 stations (figure 4.2). The three inversion modes are, 1: unconstrained atmosphere inversion, 2: joint atmosphere-ocean inversion, and 3: joint atmosphere-ocean inversion restricted to the Stephens model subset (Stephens et al., 2007). Annual results for the control case are tabulated in table 4.2, and for the sensitivity studies in tables 4.3 and 4.4.

The main result of each CO$_2$ inversion is a set of posterior fluxes and uncertainties that is most consistent with available observations, where the degree of consistency is defined by the applied transport model or models connecting data and flux spaces as well as the metric by which distances in both spaces are defined. The ability of the inversion to constrain fluxes is usually assessed based on the analysis of the parameters in flux space, such as flux uncertainty: the smaller the posterior flux uncertainty the steeper and more pronounced is the minimum of the cost function (see figure 2.8 and explanations in the accompanying text). However, it is difficult to assess the effects of biases in the data or their modeled reconstruction based on that information alone. Additional valuable information can be obtained by looking at the posterior model-data mismatch distribution, i.e. the difference between ob-
### Ocean models

<table>
<thead>
<tr>
<th>Model</th>
<th>Lat/Lon/Vert. Res.</th>
<th>Boundary forcing</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>BERN</td>
<td>(5^\circ \times 10^\circ \times 32) levels</td>
<td>Matsumoto et al. (2004)</td>
<td>Muller et al. (2006)</td>
</tr>
<tr>
<td>ECCO</td>
<td>(4^\circ \times 4.5^\circ \times 15) levels</td>
<td>NCEP</td>
<td>Stammer et al. (2004)</td>
</tr>
<tr>
<td>MIT</td>
<td>(2.8^\circ \times 2.8^\circ \times 15) levels</td>
<td>ECMWF</td>
<td>Adcroft et al. (2004)</td>
</tr>
<tr>
<td>NCAR</td>
<td>(3.6^\circ \times 1.8^\circ \times 25) levels</td>
<td>NCEP</td>
<td>Dukowicz and Smith (1994)</td>
</tr>
<tr>
<td>UL</td>
<td>(3^\circ \times 2.85^\circ \times 20) levels</td>
<td>Hellerman and Rosenstein (1983)</td>
<td>Goosse and Fichefet (1999)</td>
</tr>
<tr>
<td>PRINCE-LHS</td>
<td>(3.75^\circ \times 4.5^\circ \times 24) levels</td>
<td>da Silva et al. (1994)</td>
<td>Pacanowski and Griffies (1998), Gnanadesikan et al. (2004)</td>
</tr>
</tbody>
</table>

### Atmosphere models

<table>
<thead>
<tr>
<th>Model</th>
<th>Lat/Lon/Vert. Res.</th>
<th>Boundary forcing</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSU</td>
<td>(5^\circ \times 4.1^\circ \times 17) levels</td>
<td>Online winds</td>
<td>Denning et al. (1996)</td>
</tr>
<tr>
<td>UCB</td>
<td>(5^\circ \times 3.9^\circ \times 9) levels</td>
<td>GISS</td>
<td>Hansen et al. (1997)</td>
</tr>
<tr>
<td>UCI</td>
<td>(5^\circ \times 3.9^\circ \times 9) levels</td>
<td>GISS</td>
<td>Prather et al. (1987)</td>
</tr>
<tr>
<td>JMA</td>
<td>(2.5^\circ \times 2.5^\circ \times 32) levels</td>
<td>JMA (1997)</td>
<td>Taguchi (1996)</td>
</tr>
<tr>
<td>MATCH:CCM3</td>
<td>(2.8^\circ \times 2.8^\circ \times 28) levels</td>
<td>NCAR CCM3</td>
<td>Rasch et al. (1997)</td>
</tr>
<tr>
<td>MATCH:NCEP</td>
<td>(2.8^\circ \times 2.8^\circ \times 28) levels</td>
<td>NCEP (1990)</td>
<td>Rasch et al. (1997)</td>
</tr>
<tr>
<td>MATCH:MACCM2</td>
<td>(5.6^\circ \times 2.8^\circ \times 24) levels</td>
<td>MACCM2</td>
<td>Rasch et al. (1997)</td>
</tr>
<tr>
<td>NIES</td>
<td>(2.5^\circ \times 2.5^\circ \times 15) levels</td>
<td>ECMWF (1997)</td>
<td>Maksyutov and Inoue (2000)</td>
</tr>
<tr>
<td>NIRE</td>
<td>(2.5^\circ \times 2.5^\circ \times 15) levels</td>
<td>ECMWF (1995)</td>
<td>Taguchi (1996)</td>
</tr>
<tr>
<td>TM2</td>
<td>(7.5^\circ \times 7.2^\circ \times 9) levels</td>
<td>ECMWF (1990)</td>
<td>Bousquet et al. (1999a)</td>
</tr>
<tr>
<td>TM3</td>
<td>(5^\circ \times 4^\circ \times 19) levels</td>
<td>ECMWF (1990)</td>
<td>Heimann (1995)</td>
</tr>
</tbody>
</table>

Table 4.1 Ocean and atmosphere transport models, for which Green’s functions (footprints) were used in this study. In bold: Stephens’ model subset.
Chapter 4. Decadal trends in carbon exchange

servations and modeled responses at the stations. It should be assured that the mismatch distribution is not biased significantly, as this would imply either biased observations or insufficient model transport. The resulting bias in posterior fluxes could not be detected, because the inversion methodology assumes both unbiased observations (distributed as Gaussian) and perfect transport. The minimum of the cost function would be elevated, but this would be compensated by the iterative scaling procedure to achieve a reduced $\chi^2$ value close to one (the reduced $\chi^2$ value is proportional to the minimum cost). In figure 4.3 the model-data mismatch for all stations and months is plotted against the prior data uncertainty (the observational uncertainty). The mismatch is normalized by the prior uncertainty and color-coded for different networks. The closer a value is to zero in the scatter plot, the smaller the mismatch is relative to the prior uncertainty and the better the inversion was able to reproduce the station data accurately. Values between -1 and 1 (within the area framed by the dotted lines) indicate a mismatch that is within one standard deviation from the prior data. High absolute values mean that the inversion was not able to reproduce observations properly. The unconstrained (mode 1) and joint (mode 2) inversions are plotted separately. The mode 3 inversion is not shown, as the results look very similar to mode 2.

The prior uncertainty distributions on top of the panels move towards larger values when more stations are included. This is because the prior uncertainty depends on the number of co-located stations, i.e. stations located in the same model grid cell are de-weighted, as they would be over-represented otherwise (Gurney et al., 2003). For the 1980-89 period, the prior uncertainties reach very large values, up to 50 ppm, for the common network with 97 sites. Stations with such large uncertainties are those that did not fulfill the selection criteria, but were kept to avoid overfitting issues. These extra stations underwent a very conservative uncertainty assignment procedure, based on each station’s interannual variability during the whole Globalview period (details in section 4.3). As a result, those stations are part of the common network, but for the 1980-89 decade provide only weak constraints on the flux estimation. This avoids overfitting while the inversion is still mainly driven by observations from those stations that fulfilled the regular criteria.

From the mismatch distribution (right side of each panel) it can be seen that large mismatches occur only isolated and in both negative and positive direction. Hence, the bias for both modes and all periods is generally small and never exceeds 0.06 ppm. While the larger networks (97 or more stations) give almost identical distributions for all modes, the smaller common network (25 stations) and the individual, 32 sites network for 1980-89 cause an overfitting of the observations in the unconstrained case. This can be seen from the very narrow mismatch distribution, indicating a posterior mismatch much smaller than the prior uncertainty. When the ocean constraint is included, the situation relaxes and the mismatch distributions for all configurations exhibit a similar shape with nearly unbiased mean and a reasonable width. The common (97 sites) network exhibits a somewhat higher peak for 1980-89. This is due to the inclusion of the extra stations with their elevated uncertainties, and does
### 4.4. Results and Discussion

Period mean net flux into the atmosphere (Pg C yr\(^{-1}\)) & total uncertainty (1σ)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bor N America</td>
<td>0.54 ± 0.72</td>
<td>0.15 ± 0.34</td>
<td>0.34 ± 0.44</td>
<td>-0.01 ± 0.32</td>
<td>0.29 ± 0.51</td>
<td>-0.09 ± 0.33</td>
</tr>
<tr>
<td>Boreal Asia</td>
<td>0.04 ± 0.64</td>
<td>-0.15 ± 0.55</td>
<td>-0.25 ± 0.37</td>
<td>-0.44 ± 0.49</td>
<td>-0.24 ± 0.63</td>
<td>-0.37 ± 0.50</td>
</tr>
<tr>
<td>Temp N America</td>
<td>-1.87 ± 0.69</td>
<td>-1.12 ± 0.39</td>
<td>-1.73 ± 0.89</td>
<td>-0.58 ± 0.37</td>
<td>-2.17 ± 1.17</td>
<td>-0.86 ± 0.39</td>
</tr>
<tr>
<td>Europe</td>
<td>-0.59 ± 0.55</td>
<td>-1.00 ± 0.24</td>
<td>-0.32 ± 0.59</td>
<td>-0.74 ± 0.29</td>
<td>-0.33 ± 0.64</td>
<td>-1.14 ± 0.28</td>
</tr>
<tr>
<td>Temp Asia</td>
<td>-0.53 ± 1.10</td>
<td>0.00 ± 0.38</td>
<td>-0.70 ± 0.53</td>
<td>-0.02 ± 0.36</td>
<td>-0.63 ± 0.64</td>
<td>-0.07 ± 0.36</td>
</tr>
<tr>
<td>Trop America</td>
<td>2.25 ± 1.58</td>
<td>-0.18 ± 0.54</td>
<td>3.51 ± 1.15</td>
<td>-0.05 ± 0.58</td>
<td>3.32 ± 1.42</td>
<td>0.21 ± 0.60</td>
</tr>
<tr>
<td>Northern Africa</td>
<td>-0.09 ± 1.81</td>
<td>0.02 ± 0.35</td>
<td>-0.50 ± 1.31</td>
<td>-0.10 ± 0.36</td>
<td>-0.50 ± 1.88</td>
<td>-0.06 ± 0.39</td>
</tr>
<tr>
<td>Trop Asia</td>
<td>-0.25 ± 0.67</td>
<td>0.51 ± 0.30</td>
<td>0.07 ± 0.36</td>
<td>0.37 ± 0.32</td>
<td>0.03 ± 0.25</td>
<td>0.44 ± 0.31</td>
</tr>
<tr>
<td>South America</td>
<td>-1.21 ± 1.14</td>
<td>0.23 ± 0.33</td>
<td>-1.89 ± 1.18</td>
<td>0.12 ± 0.33</td>
<td>-2.19 ± 1.22</td>
<td>-0.10 ± 0.33</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>1.38 ± 1.05</td>
<td>0.64 ± 0.58</td>
<td>0.02 ± 0.97</td>
<td>-0.12 ± 0.56</td>
<td>1.07 ± 1.06</td>
<td>0.04 ± 0.62</td>
</tr>
<tr>
<td>Australasia</td>
<td>-0.17 ± 0.17</td>
<td>-0.17 ± 0.11</td>
<td>0.21 ± 0.26</td>
<td>-0.09 ± 0.09</td>
<td>-0.07 ± 0.24</td>
<td>-0.34 ± 0.12</td>
</tr>
<tr>
<td>Northern Ocean</td>
<td>-0.18 ± 0.09</td>
<td>n/a</td>
<td>-0.18 ± 0.09</td>
<td>n/a</td>
<td>-0.18 ± 0.09</td>
<td>n/a</td>
</tr>
<tr>
<td>N Pacific</td>
<td>-0.43 ± 0.04</td>
<td>n/a</td>
<td>-0.43 ± 0.04</td>
<td>n/a</td>
<td>-0.44 ± 0.04</td>
<td>n/a</td>
</tr>
<tr>
<td>N Atlantic</td>
<td>-0.33 ± 0.04</td>
<td>n/a</td>
<td>-0.34 ± 0.04</td>
<td>n/a</td>
<td>-0.35 ± 0.04</td>
<td>n/a</td>
</tr>
<tr>
<td>W Pacific</td>
<td>0.09 ± 0.05</td>
<td>n/a</td>
<td>0.09 ± 0.05</td>
<td>n/a</td>
<td>0.09 ± 0.05</td>
<td>n/a</td>
</tr>
<tr>
<td>E Pacific</td>
<td>0.24 ± 0.07</td>
<td>n/a</td>
<td>0.23 ± 0.07</td>
<td>n/a</td>
<td>0.22 ± 0.07</td>
<td>n/a</td>
</tr>
<tr>
<td>Trop Atlantic</td>
<td>0.16 ± 0.03</td>
<td>n/a</td>
<td>0.15 ± 0.03</td>
<td>n/a</td>
<td>0.15 ± 0.03</td>
<td>n/a</td>
</tr>
<tr>
<td>Trop Indian</td>
<td>0.10 ± 0.02</td>
<td>n/a</td>
<td>0.10 ± 0.02</td>
<td>n/a</td>
<td>0.10 ± 0.02</td>
<td>n/a</td>
</tr>
<tr>
<td>S Pacific</td>
<td>-0.45 ± 0.09</td>
<td>n/a</td>
<td>-0.46 ± 0.09</td>
<td>n/a</td>
<td>-0.46 ± 0.09</td>
<td>n/a</td>
</tr>
<tr>
<td>S Atlantic</td>
<td>-0.19 ± 0.02</td>
<td>n/a</td>
<td>-0.19 ± 0.02</td>
<td>n/a</td>
<td>-0.20 ± 0.02</td>
<td>n/a</td>
</tr>
<tr>
<td>S Indian</td>
<td>-0.48 ± 0.06</td>
<td>n/a</td>
<td>-0.49 ± 0.08</td>
<td>n/a</td>
<td>-0.49 ± 0.08</td>
<td>n/a</td>
</tr>
<tr>
<td>Southern Ocean</td>
<td>-0.27 ± 0.10</td>
<td>n/a</td>
<td>-0.28 ± 0.10</td>
<td>n/a</td>
<td>-0.31 ± 0.10</td>
<td>n/a</td>
</tr>
<tr>
<td>Northern Land</td>
<td>-2.41 ± 0.59</td>
<td>-2.10 ± 0.63</td>
<td>-2.65 ± 0.43</td>
<td>-1.80 ± 0.60</td>
<td>-3.08 ± 0.45</td>
<td>-2.53 ± 0.58</td>
</tr>
<tr>
<td>Tropical Land</td>
<td>1.92 ± 1.11</td>
<td>0.34 ± 0.77</td>
<td>3.08 ± 1.01</td>
<td>0.22 ± 0.82</td>
<td>2.85 ± 1.24</td>
<td>0.59 ± 0.84</td>
</tr>
<tr>
<td>Tropical Africa+Asia</td>
<td>-0.34 ± 1.63</td>
<td>0.53 ± n/a</td>
<td>-0.43 ± 1.52</td>
<td>0.27 ± n/a</td>
<td>-0.47 ± 1.79</td>
<td>0.38 ± n/a</td>
</tr>
<tr>
<td>Southern Land</td>
<td>0.00 ± 0.71</td>
<td>0.71 ± 0.58</td>
<td>-1.67 ± 0.70</td>
<td>-0.09 ± 0.60</td>
<td>-1.20 ± 0.93</td>
<td>-0.40 ± 0.60</td>
</tr>
<tr>
<td>Tropical-Southern Land</td>
<td>1.92 ± 0.61</td>
<td>1.05 ± n/a</td>
<td>1.41 ± 0.44</td>
<td>0.13 ± n/a</td>
<td>1.65 ± 0.47</td>
<td>0.19 ± n/a</td>
</tr>
<tr>
<td>Northern Ocean</td>
<td>-0.94 ± 0.08</td>
<td>n/a</td>
<td>-0.95 ± 0.08</td>
<td>n/a</td>
<td>-0.97 ± 0.08</td>
<td>n/a</td>
</tr>
<tr>
<td>Tropical Ocean</td>
<td>0.59 ± 0.10</td>
<td>n/a</td>
<td>0.57 ± 0.10</td>
<td>n/a</td>
<td>0.56 ± 0.10</td>
<td>n/a</td>
</tr>
<tr>
<td>Ocean in South</td>
<td>-1.39 ± 0.15</td>
<td>n/a</td>
<td>-1.42 ± 0.15</td>
<td>n/a</td>
<td>-1.46 ± 0.16</td>
<td>n/a</td>
</tr>
<tr>
<td>Global Land</td>
<td>-0.49 ± 0.27</td>
<td>-1.10 ± 0.48</td>
<td>-1.24 ± 0.25</td>
<td>-1.66 ± 0.51</td>
<td>-1.42 ± 0.27</td>
<td>-2.34 ± 0.56</td>
</tr>
<tr>
<td>Global Ocean</td>
<td>-1.74 ± 0.24</td>
<td>n/a</td>
<td>-1.80 ± 0.25</td>
<td>n/a</td>
<td>-1.88 ± 0.25</td>
<td>n/a</td>
</tr>
<tr>
<td>Global Total</td>
<td>-2.23 ± 0.13</td>
<td>n/a</td>
<td>-3.03 ± 0.04</td>
<td>n/a</td>
<td>-3.30 ± 0.10</td>
<td>n/a</td>
</tr>
</tbody>
</table>

**Table 4.2** Annual posterior fluxes for all 22 regions in the inversion as well as selected aggregated regions. Control results (mode 3, common network) are compared to the study of Gurney and Eckels (2011) (GE). Bold numbers indicate 90% statistical significance for the flux sign. Highlighted cells contain fluxes that differ by more than 0.5 Pg C yr\(^{-1}\) from the individual network inversion (see table in appendix for more details).

not indicate overfitting. In many other atmospheric inversions with monthly flux resolution that are based on the Transcom regions (e.g. Gurney et al. (2008); Gurney and Eckels (2011)), the overfitting issue does not show up. For example, Gurney et al. (2008) applied eight different networks to their interannual inversion without facing that issue. This is because they used monthly prior fluxes for all land regions, which resolves the problem in the same
way the inclusion of our ocean constraint does. However, the addition of prior information to an otherwise overfitted inversion can enhance the dominance of the prior over the data. To avoid this, we decided to exclude the 25 sites network and use the 97 sites network as common network instead. In addition, we will treat the results from the unconstrained 1980-89 inversion on the 32 sites individual network with caution, i.e. results will be reported but not analyzed in depth.

4.4.1 Oceanic trends and consequences on land

In the joint inversion modes 2 and 3 the regional ocean fluxes are mainly determined by the ocean inversion. Due to the tight uncertainty estimates the atmospheric CO$_2$ constraint has only little influence on the joint estimate. In particular, the longterm global ocean sink is known with high confidence, due to the vast amount of DIC data available for the ocean inversion. Oceanic fluxes are a composite of a preindustrial and an anthropogenic distribution. Preindustrial fluxes were corrected for the riverine carbon loop, with the correction being distributed regionally according to Jacobson et al. (2007a) and summing up globally to 0.45 Pg C yr$^{-1}$. The correction is constant over time, i.e. no assumptions were made about interannual variations in the river carbon loads. Anthropogenic fluxes were scaled to individual years according to equation 4.1, before both contributions were summed to yield the net ocean flux that enters the joint inversion. The anthropogenic scaling function $\phi(t)$ (plotted in figure 4.5) depends on the atmospheric CO$_2$ perturbation as well as its history (see Mikaloff Fletcher et al. (2006); Gloor et al. (2003) for more details).

**Figure 4.5** Perturbation of atmospheric CO$_2$ compared to a preindustrial value of 280 ppm is shown as dashed line. The solid curve represents the function $\phi(t)$ used to scale the anthropogenic ocean flux, with year 2000 as a reference. See text for details.
The net oceanic fluxes therefore include the riverine carbon correction as well as an interannual component linked to atmospheric CO$_2$. They do not account for, however, any other interannual variations, such as a direct oceanic response to El Nino events. An indirect El Nino response may be seen via its imprint on atmospheric CO$_2$, for example for the strong El Nino in 1997/8. Interannual changes in ocean circulation are also not considered, as the transport models underlying the ocean inversion were run with invariant circulation. In consequence, the curvature of the oceanic sink depicted in figure 4.1 is determined directly by the scaling function in figure 4.5.

To compare our joint ocean fluxes with other studies, it is often necessary to subtract the riverine correction again, as many studies report the anthropogenic flux rather than the net flux. For the global flux this can be done by increasing our estimates by 0.45 Pg C yr$^{-1}$, because the non-corrected preindustrial flux is very close to zero.

When averaged over the inversion periods, the global oceans take up -1.74 ($\pm$0.24) Pg C yr$^{-1}$, -1.80 ($\pm$0.25) Pg C yr$^{-1}$, and -1.88 ($\pm$0.25) Pg C yr$^{-1}$ for 1980-89, 1990-99, and 2000-08, respectively. This translates into a trend of -0.007 Pg C yr$^{-2}$, which is distributed evenly across oceanic regions, because the anthropogenic scaling was applied spatially homogeneous. The negative trend implies that the global oceans sequestered an extra amount of 0.2 Pg C since 1980. This estimate is in very good agreement with recent model-based studies (e.g. Raupach (2011); Le Quere et al. (2009); Sarmiento et al. (2010) and references therein). Interannual variations are generally more pronounced in those models, as they are usually forced by re-analyzed and interannually varying wind, freshwater and heat fields. However, oceanic interannual variability is still small compared to that of the land-atmosphere CO$_2$ exchange (figure 1 in Sarmiento et al. (2010)). This makes us confident that, although we do not capture the full interannual variations in sea-air CO$_2$ fluxes, we have good estimates for their decadal mean as well as trend. In conjunction with the well known global sink deduced from fossil emissions and atmospheric growth rate, the decadal mean and trend of the land sink can be inferred with high confidence, too.

At the regional level our oceanic trends are less reliable, in particular for regions sensitive to climate change effects. The most prominent example is the Southern Ocean sink, for which recent studies found indications of saturation in the last few decades. Le Quere et al. (2007) found a weakening of the Southern Ocean sink of 0.2 Pg C yr$^{-1}$ between 1979 and 2004, relative to the expected sink evolution if driven by atmospheric CO$_2$ perturbation only. They attribute this to anthropogenically induced changes in the wind fields. For the same time period, Lovenduski et al. (2008) found a relative weakening of 0.1 Pg C yr$^{-1}$, when they compared model hindcast simulations using re-analyzed forcings with simulations with fixed physical climate. Note that both studies did not find a decrease of the Southern Ocean sink, but an increase that is weaker than expected, though still an increase.
We tested the impact that such a saturation of the Southern Ocean sink would have on our joint inversion by setting the Southern Ocean flux to a constant value throughout all inversion periods, i.e. assuming a complete saturation of the sink. We chose zero as constant value, which represents a drastic change in Southern Ocean flux between 0.2 and 0.3 Pg C yr\(^{-1}\) (figure 4.6) for the three decades. This must be compensated over land, as the remaining oceanic fluxes stay untouched. It turns out that the land response is distributed over four regions, the tropical as well as the southern parts of America and Africa. All other regions show no significant response. The South American sink reacts the most sensitive and increases by more than 1 Pg C yr\(^{-1}\), the Northern African region takes up an additional 0.2 to 0.3 Pg C yr\(^{-1}\), depending on the inversion mode. The Tropical American and Southern African regions release approximately 0.4 and 0.7 Pg C yr\(^{-1}\) more. Hence, the absence of the Southern Ocean sink leads to strong responses in these regions, particularly in South America. The missing sink is not just shared among the regions, but causes both additional outgassing and uptake that partially compensate for each other. The general pattern is that already existing sources (Tropical America, Southern Africa) and sinks (Northern Africa, South America) are enhanced further, thus strengthening the source/sink dipole on both continents. In their study, Sarmiento et al. (2010) estimate the net tropical land flux to be close to zero for the 1990-99 decade based on a bottom-up approach. They find \(-0.1 (\pm 0.7)\) Pg C yr\(^{-1}\) for Tropical America and \(0.0 (\pm 0.9)\) Pg C yr\(^{-1}\) for the remaining tropical land in Africa and Asia. Our control inversion assigns already a large carbon source over Tropical America and a moderate sink over Northern Africa (see table 4.2). This discrepancy gets more severe when this Southern Ocean scenario is considered.

The inverse relationship between the Southern Ocean sink strength and the source-sink dipole strength on the African and American continents was already apparent in previous inversion studies using the Transcom models. For example, the atmospheric inversion of Gurney et al. (2004) estimates a Southern Ocean sink of \(-0.55\) Pg C yr\(^{-1}\) for the 1992-96 period and relatively weak dipole magnitudes of \(1.3\) and \(1.0\) Pg C yr\(^{-1}\) for Africa and America. Jacobson et al. (2007b), on the other hand, estimate a smaller Southern Ocean sink of \(-0.15\) Pg C yr\(^{-1}\) (same period) accompanied by strong dipoles of \(3.4\) and \(4.0\) Pg C yr\(^{-1}\) for Africa and America.

To summarize, it should be kept in mind that regional oceanic fluxes account for atmospheric CO\(_2\) evolution, but not for any other interannual variations, nor for longer term trends caused by anthropogenic activities. For decadal mean fluxes and trends, the underestimated interannual variability is not expected to cause large biases, in particular for the well constrained global ocean sink strength and trend. Erroneous longterm trends at the regional level can, however, have major impacts on some land fluxes. Our Southern Ocean experiment is an example, though quite extreme due to the drastic changes applied. We think that the inclusion of our scaled ocean prior is still preferable to an ocean prior based on ocean biogeochemistry models, because it is constrained by observations more directly, and it contains further
4.4. Results and Discussion

![Graph showing changes in annual fluxes in response to a prescribed Southern Ocean flux of zero.](image)

**Figure 4.6** Change in annual fluxes in response to a prescribed Southern Ocean flux of zero. Changes are based on the difference between results from the inversions mode 2 (ocean constraint) as well as mode 3 (ocean and model constraints) and the respective inversions where the Southern Ocean flux is prescribed as zero. The increasing positive flux change over the periods is caused by the fact that the Southern Ocean sink is estimated to increase over time in the normal inversions.

Information about errors and correlations that are utilized in the joint inversion. In the following subsections we will study the effect our ocean constraint has on the estimation of land-atmosphere CO$_2$ exchange.

### 4.4.2 Global sink and latitudinal partitioning

The global sink for atmospheric CO$_2$ over ocean and land is known to a high degree of certainty, because it is determined from well-known fossil fuel emissions and observed growth rate of CO$_2$ in the atmosphere. 2.23 (±0.13) Pg C yr$^{-1}$ were sequestered from the atmosphere on average between 1980 and 1989, followed by an increase to 3.03 (±0.04) Pg C yr$^{-1}$ in the 1990s, and 3.30 (±0.10) Pg C yr$^{-1}$ over 2000-2008.

From the 1.07 Pg C yr$^{-1}$ decadal mean difference, only 0.14 Pg C yr$^{-1}$ are contributed by the global oceans. The remaining 0.93 Pg C yr$^{-1}$ are due to increased uptake by the land...
biosphere, in agreement with previous studies (Sarmiento et al., 2010; Le Quere et al., 2009). Gurney and Eckels (2011) find a larger increase of 1.24 Pg C yr\(^{-1}\), which is a substantial deviation considered that they use the same Transcom setup and similar global growth and fossil emissions data. The reason is their loose ocean prior and the resulting unrealistic ocean sink trend (see section 4.4.3 for further discussion). Consequently, the role of the land biosphere in global carbon uptake has grown bigger over the last three decades: the land uptake portion doubled from 22% for 1980-89 to 43% for 2000-08 (thick bars in figure 4.7; see also figure 4.1). Most of this doubling, 41%, occurred already until the 1990s, with only little additional increase thereafter. This is consistent with the finding of Sarmiento et al. (2010) of a sudden increase in land uptake after 1988/89, with maximum uptake between 1991 and 1993, and followed by a relatively constant - though larger than before 1988/89 - uptake until 2006 (the end of their period of study). They attribute the maximum uptake to the Pinatubo eruption in 1991, but also point out that the land started shifting to a higher overall uptake already before, i.e. in the second half of the 1980s.

![Figure 4.7](image)

**Figure 4.7** Partitioning of the global carbon sink in the control inversion. For each period three levels are considered: 1) partitioning of global sink into land and ocean uptake, 2) split of the global land uptake into fluxes from Northern (extra-tropical) and Tropical plus Southern Land (TSL) regions, and 3) individual contributions from Tropical and Southern land regions to the TSL flux. All fluxes are normalized by the period-mean global carbon sink, which is added as number below the global sink bar. Errorbars represent normalized 1\(\sigma\) total uncertainty, i.e. the quadrature of within-uncertainty and model spread.

Their absolute values are -0.27 Pg C yr\(^{-1}\) averaged before 1988 and -1.15 Pg C yr\(^{-1}\) averaged over 1989 and 2007. This jump is also visible in our results (table 4.2) that give -0.49 (±0.27) Pg C yr\(^{-1}\) for 1980-89, and then -1.24 (±0.25) Pg C yr\(^{-1}\) and -1.42 (±0.27) Pg C yr\(^{-1}\) for 1990-99 and 2000-08, respectively. By contrast, GE do not find this pattern: they estimate the land sink to increase from -1.10 (±0.48) to -1.66 (±0.51) and -2.34 (±0.56), hence only moderately in the 90s and then much stronger in the 2000s, when Sarmiento et al. (2010) do
not see a significant change.

A significant increase in overall land uptake towards the end of the 1980s is also supported by land biosphere model simulations (Le Quere et al., 2009). By contrast, Gurney and Eckels (2011) find a much more linear increase in land uptake over the past three decades: they estimate the land sink to increase from -1.10 (±0.48) to -1.66 (±0.51) and -2.34 (±0.56), hence only moderately in the 90s and then much stronger in the 2000s, when we and Sarmiento et al. (2010) do not see a significant change.

For all three decades the global land sink is composited of a sink over Northern Hemisphere (extra-tropical) land and a source from Tropical and Southern land (TSL) regions. The TSL source is in turn the result of a Tropical source and the flux from Southern Hemisphere (extra-tropical) land. The latitudinal breakdown of relative contributions to the global sink is shown in figure 4.7. The major driver for the abrupt increase in land uptake between the 1980s and 1990s is a reduction of TSL source from 86% to 47% of the global sink. The sink contribution from Northern Land declined at the same time from 108% to 87%, thus partially compensated the TSL effect, which otherwise would have led to an even stronger overall land uptake. Hence, while the absolute Northern Land sink steadily increased over the decades (figure 4.8), its relative contribution to the overall land sink decreased after 1980-89 and recovered only partially to 93% in 2000-08.

Partitioning the TSL region into Tropical and Southern land reveals that in the 1980s the TSL source can be attributed completely to the Tropics. There was no carbon flux from or into the Southern land biosphere. In the 1990s this picture changed dramatically: Southern land regions now took up 1.67 (±0.70) Pg C yr⁻¹, corresponding to 55% of the global sink. At the same time the Tropical source increased, though not enough to compensate for the Southern land uptake. Between 2000 and 2008 the situation was similar, with a somewhat smaller uptake fraction from Southern land accompanied by a reduction in the Tropical source fraction of equal magnitude. The partitioning in 2000-08 is generally similar to the 1990-99 picture. The land portion of the global sink did not change much, and so did the TSL source fraction. However, there was a slight redistribution of uptake from Southern to Northern land, both in absolute magnitude (figure 4.8) and relative contribution.

### 4.4.3 Decadal trends on land: global picture

We now turn from the consideration of relative contributions to the discussion of absolute flux changes and trends based on linear regressions on the decadal mean fluxes. Results for land regions and aggregates are presented in figure 4.8, whereas figure 4.9 shows all regions including the ocean and according aggregates. Numerical values are given in tables 4.2 for the control case and in the appendix tables 4.3, 4.4 for all inversion modes. Here we
Figure 4.8  Decadal flux changes in Pg C yr⁻¹ for land regions (world map) as well as latitudinal aggregates (right panel). The 1980-2008 average was subtracted to show relative changes. Trends for the 1980-2008 period were calculated by a regression and are given as numbers in each inset. In addition to the control case (green), results are shown from the unconstrained inversion on the common network (blue), and the inversion with ocean and model constraints on individual networks (dashed green). Estimates of Gurney and Eckels (2011) are added in red. Connecting lines are only shown for clarity, they do not represent interannual variation.
4.4. Results and Discussion

give a summary of decadal trends by latitude and compare them to a recent study of Gurney and Eckels (2011) (henceforth, GE). A more detailed and regionally resolved picture is then given in the following sections.

As mentioned above, the relative contribution from Northern land to the global land sink decreases. However, the absolute Northern land sink continuously increases throughout the three decades, exhibiting a linear trend of \(-0.034\) Pg C yr\(^{-2}\). GE estimate a slightly smaller trend of \(-0.022\) Pg C yr\(^{-2}\). This number is based on a linear regression on their decadal mean fluxes in the same manner like our estimates, in order to make the two directly comparable. Trends directly reported by them are based on a regression on annual fluxes, though with generally similar results. Boreal areas contribute 79% to the Northern land trend, mainly driven by an intensification of growing season uptake.

Southern land regions switched from being carbon neutral in the 1980s to a large sink of \(-1.67 (\pm 0.70)\) Pg C yr\(^{-1}\) in the 1990s, followed by a slightly smaller sink in the 2000s. They exhibit an overall trend of \(-0.062\) Pg C yr\(^{-2}\), by far the strongest trend of the three main latitudinal bands. The trend found by GE is of similar magnitude, \(-0.057\) Pg C yr\(^{-2}\). The temperate regions in South America as well as Southern Africa dominate this trend, though their individual contributions vary strongly for different decades. The surprisingly important role of Southern Africa is a finding consistent with GE (cf. also review of Raupach (2011)), who relate this to decadal changes in rainfall patterns.

In this study, the Tropics are found to be a net CO\(_2\) source to the atmosphere throughout all decades. The overall trend is \(0.048\) Pg C yr\(^{-2}\), indicating a pantropical carbon release that increases at a markedly higher rate than estimated by GE \((0.013\) Pg C yr\(^{-2}\)). Even when a pantropical deforestation estimate of \(1.5\) Pg C yr\(^{-1}\) (Le Quere et al., 2009; Canadell et al., 2007) is subtracted, fluxes from all decades stay positive, implying a net carbon release from undisturbed tropical forest. This disagrees with most bottom-up studies (cf. discussion in chapter 3), which assign a small vegetational sink to these areas. However, an enhanced tropical source in conjunction with an enhanced southern land sink is an expression of underlying spatial correlations typical for atmospheric inversions. The dipole character results from the fact that their sum is well-known, but their individual fluxes are poorly constrained due to the lack of observations in these regions. We find a rather strong tropical-southern land dipole, because we do not use land priors, which otherwise would act as rectifier. In their joint inversion, Jacobson et al. (2007a) find an even stronger dipole effect, driven by their rejection of land priors. By contrast, the Transcom-based atmospheric inversions (e.g. Gurney et al. (2004); Gurney and Eckels (2011); Baker et al. (2006)) make use of land priors, resulting in a much smaller dipole effect. Our ocean prior also regularizes tropical fluxes, though weaker and only indirectly through land-ocean coupling, as can be seen by comparison with our unconstrained (mode 1) inversion. However, it imposes different decadal flux changes in some regions than the Transcom land prior, as will be discussed below in more
Figure 4.9 Decadal flux changes in Pg C yr\(^{-1}\) for all 11 land and 11 ocean regions as well as latitudinal aggregates. The 1980-2008 average was subtracted to show relative changes. Trends for the 1980-2008 period were calculated by a regression and are given as numbers in each inset. In addition to the control case (blue), results are shown from the unconstrained inversion on the common (gray) and individual (dashed gray) network, and the inversion with ocean and model constraints on individual networks (dashed blue). Estimates of Gurney and Eckels (2011) are added in green. They did not report ocean fluxes, therefore the trends are set to NaN for all oceanic regions as well as aggregates involving oceanic regions. Results from the unconstrained inversion on individual networks are excluded from analysis due to instability issues in the 1980-89 period (see text for further information).
4.4. Results and Discussion

detail.

Together, the flux trends from the latitudinal aggregates underlie the trend for overall land uptake of \(-0.048\) Pg C yr\(^{-2}\), accounting for 87% of the trend in global uptake \((-0.055\) Pg C yr\(^{-2}\)). GE find an even stronger trend in overall land uptake of \(-0.064\) Pg C yr\(^{-2}\). They did not report their estimates for oceanic and global uptake, but since they also constrained global uptake based on the observed global growth rate and fossil fuel emissions from the same datasets, it is very likely they obtained a very similar trend in global uptake. This would imply a positive trend of \(0.009\) Pg C yr\(^{-2}\) in their global oceanic sink, corresponding to \(0.26\) Pg C yr\(^{-1}\) that the oceans took up less over the 29 years from 1980 to 2008. When using their reported value of \(-0.057\) Pg C yr\(^{-2}\) for the global land trend, which is based on a regression over all years within 1980-2008, the effect becomes smaller, but still implies a reduction in global ocean sink strength of \(0.06\) Pg C yr\(^{-1}\). There are indications that some regions in the ocean do not increase in sink strength any more, e.g. the Southern Ocean (studies of Le Quere et al. (2007); Lovenduski et al. (2008) and discussion in section 4.4.1). However, there is no reason to believe that, e.g., the Southern Ocean sink saturation turns into a sink reduction. Moreover, on a global scale indications are strong that the ocean sink still increases in response to continued atmospheric CO\(_2\) accumulation (e.g. Raupach (2011); Le Quere et al. (2009)). Hence, from the current viewpoint a positive trend in global ocean uptake appears unrealistic.

In a recent ecosystem model (ORCHIDEE) study, Piao et al. (2009) estimate the global land sink to have intensified from \(-0.5\) Pg C yr\(^{-1}\) in the 1980s to \(-1.0\) Pg C yr\(^{-1}\) in the 1990s, in good agreement with our results \((-0.49\) (±0.27) Pg C yr\(^{-1}\) and \(-1.24\) (±0.25) Pg C yr\(^{-1}\), respectively). This suggests that the ocean constraint leads to realistic land uptake estimates when combined with total uptake constraints from atmospheric CO\(_2\) accumulation and growth rate. GE were not able to reproduce these estimates, in particular there land sink in the 1980s appears somewhat large \((-1.10\) (±0.48) Pg C yr\(^{-1}\)) and implies a very small ocean sink during that decade.

Our ocean constraint avoids the unrealistic positive global ocean trend. The global oceans exhibit a small negative trend of \(-0.007\) Pg C yr\(^{-2}\), yielding an increase in uptake of \(0.20\) Pg C yr\(^{-1}\) for the 1980-2008 period. As a result, the increase in land uptake is not as large as found by GE (0.31 Pg C yr\(^{-1}\) smaller change between mid-1980s and mid-2000s). Nevertheless, we find slightly stronger trends than GE in both Northern and Southern land regions (figure 4.8). But these are over-compensated by a large increase in Tropical land source.

Latitudinal trends are generally robust against choosing individual or common networks. This is an anticipated result, since previous studies (Gurney et al., 2008; Roedenbeck et al., 2003b) have shown that temporal flux variations are less sensitive to the choice of network than absolute values and spatial flux distribution. Interannual variability as well as longterm
trends seem the most robust features, except for a few regions mostly located in the tropics (Gurney et al., 2008). In this study, the tropics and also the southern land exhibit slight systematic shifts towards more pronounced trends when the individual setup is used. These are mostly due to deviations in the 1980s, which are caused by 1) larger differences in network size (i.e. number of stations) compared to the more recent decades, and 2) the small size of the 1980s individual network, leading to the aforementioned instabilities (the number of stations approaches the number of flux parameters). As will be seen in the following sections, trends at the regional level remain insensitive to network choice for most regions, but there are important exceptions, such as Europe, where the deviations among networks dominate over those from different inversion modes.

Latitudinal trends agree in sign for all inversion modes. The positive tropical trend as well as the negative trends in the northern and southern land regions weaken in the control (mode 3) inversion compared to mode 2, owing to intensified ventilation in the models selected for mode 3 (discussions in chapter 3 and by Stephens et al. (2007)). Even the unconstrained (mode 1) inversion produces similar, though steeper (except for southern land), trends. Period-mean, absolute fluxes deviate, however, largely between modes 1 and 2 (table 4.3).

4.4.4 Northern land

The negative flux trend (increasing uptake) for northern land is driven primarily by boreal areas. Decadal mean fluxes from boreal land exhibit a negative trend of -0.027 Pg C yr\(^{-2}\), resulting from an intensification of growing season net flux (GSNF; -0.015 Pg C yr\(^{-2}\)) combined with a decline in dormant season net flux (DSNF; -0.012 Pg C yr\(^{-2}\)), and only little variations in seasonal amplitude. By comparison, Gurney and Eckels (2011) (GE) find an increase in amplitude in absence of a significant DSNF trend. GSNF and DSNF and seasonal amplitude are defined following GE. Briefly, the GSNF is obtained by summing monthly fluxes for months exhibiting negative fluxes (uptake from the atmosphere), and the DSNF by summing monthly fluxes for months exhibiting positive fluxes (release to the atmosphere). Trends in GSNF and DSNF are calculated the same way like annual net flux trends.

Boreal America and Asia show very similar trends of -0.013 and -0.014 Pg C yr\(^{-2}\), in agreement with GE and also Chevallier et al. (2010), who find increasing boreal Asia uptake (-0.017 Pg C yr\(^{-2}\)) using an atmospheric inversion with interannually varying transport. However, we find the seasonal drivers for these trends to be different for the two regions, contrasting to GE, who attribute the trends in both regions predominantly to an intensifying GSNF. For boreal America we also find GSNF intensification (-0.015 Pg C yr\(^{-2}\), see also figure 4.10a) to be the main driver. Boreal Asia, on the other hand, exhibits a weakening of GSNF in combination with a decline in DSNF and accordingly a decrease in seasonal amplitude. Therefore, the
4.4. Results and Discussion

overall increase in boreal Asia uptake is due to less outgassing during the dormant season and does not result from intensified GSNF. In fact, when averaging over all Transcom models (mode 2), the weakening of GSNF even leads to a slightly positive trend in boreal Asia uptake (decrease in uptake, see figure 4.9). This positive trend vanishes when switching to the individual network setup. The flux estimate for boreal Asia is generally more sensitive to network choice than for boreal America, because there are no stations located directly inside boreal Asia (all networks).

We estimate the temperate northern land regions to take up an increasing amount of carbon over the decades (-0.006 Pg C yr\(^{-2}\)), whereas GE find a slight weakening (0.005 Pg C yr\(^{-2}\)). Although both trends are small and not significant, they point towards one major difference between the results here and those of GE: we find a decreasing European (0.014 Pg C yr\(^{-2}\)) and an increasing temperate north American sink (-0.015 Pg C yr\(^{-2}\)), while GE estimate trends with opposite signs (Europe -0.007 Pg C yr\(^{-2}\) and temperate north America 0.014 Pg C yr\(^{-2}\)). The discrepancy becomes larger for mode 2, where the temperate north American sink intensification becomes about twice as large. In mode 1 (unconstrained) the sink intensification reaches -0.05 Pg C yr\(^{-2}\). The weakening of the European sink, on the other hand, is extremely robust against mode changes. It is caused by an increased DSNF while the GSNF remains fairly constant over the decades, thus increasing the seasonal amplitude asymmetrically. This suggests elevated winter respiration and only little changes in NPP (net primary production) during the growing season. The latter may hide countervailing effects in growing-season gross productivity and heterotrophic respiration, which cannot be resolved within the scope of this study. The increasing temperate north American sink is also dominated by a decline of the DSNF rather than changes in the GSNF.

Temperate northern land trends in America, Europe and Asia are very sensitive to the choice of networks: the negative north American and positive European trends intensify remarkably when switching from the common network to individual networks. The temperate Asian sink trend does not change much, though the 1990s exhibit the weakest decadal uptake when using individual networks, opposite to the common setup that assigns the strongest decadal uptake during the 1990s. For both network setups mode changes are generally of minor importance in Europe and temperate Asia. An exception is the estimation of a huge source for temperate Asia in the 1980s when using the individual network with 32 sites in an unconstrained inversion, due to the lack of sites there. The two closest stations are located in the western tropical Indian and Pacific oceans, making it for the inversion impossible to distinguish between the temperate and tropical parts of Asia. This leads to a huge compensating sink in tropical Asia and demonstrates once more that this particular inversion setup should not be over-emphasized, since it is not able to resolve 22 regions properly. The temperate north American trend is subject to similar sensitivity to the choice of networks and modes.

The increase in combined northern temperate land uptake is robustly estimated across net-
works and modes, though is distribution between Europe and America depends sensitively on the choice of network. This indicates that fluxes from these regions are primarily constrained by atmospheric CO$_2$ and that further inclusion of the ocean prior as well as the se-
lection of transport models have only minor influence. While this is an anticipated result, the network-sensitive redistribution of the American and European fluxes is unexpected, in particular between the 1990s and the 2000s, when differences in networks are relatively small (the common network is a subset of the individual networks from the 1990s and 2000s). Yet, the additional stations - located in the northern USA, at the southern and east coasts of the Mediterranean sea, at the black sea and in northern Africa, cf. figure 4.2 - are responsible for the strong weakening of the European carbon sink and even estimate a net carbon release of $0.63 \pm 0.61$ Pg C yr$^{-1}$ from Europe during 2000-08.

Absolute fluxes indicate that the northern land has removed about $-2.7$ Pg C yr$^{-1}$ on average over the three decades examined here. This estimate is relatively high compared to other studies, e.g. GE estimated $-2$ Pg C yr$^{-1}$ averaged over the same time period, Ciais et al. (2010) found $-1.7 \pm 0.9$ Pg C yr$^{-1}$ over the period 2000-04. The estimate of Ciais et al. is based on four different atmospheric inversion systems and several bottom-up estimates. We obtain a stronger northern land sink, because we do not use prior information on land fluxes in our inversion. As has been shown in chapter 3 and in previous studies (e.g. Jacobson et al. (2007a) this leads to stronger northern land uptake in concert with stronger tropical land outgassing. The effect can also be seen here by comparing our inversion modes 1 and 2. When comparing our results with atmospheric studies using land priors, it should be kept in mind that our aim was primarily to estimate decadal flux trends rather than absolute fluxes. Our trend estimates are based on atmospheric CO$_2$ and oceanic DIC and pCO$_2$ data only, and thus isolate their information content. It is generally remarkable that our trend estimates for all regions (not only latitudinal bands) approach those of GE (who used land priors) when we include the ocean prior. Hence, it is possible to obtain realistic land flux trends only through atmospheric and oceanic carbon measurements, without assuming decadal flux changes over land. The latter are hard to find in the literature, because inventory and flux tower studies usually cover only the 2000s and part of the 1990s.

Potter et al. (2003) isolated flux trends for the north American continent based on satellite measurements (AVHRR) and ecosystem modeling (CASA model). They find a carbon sink varying between $-0.2$ and $-0.3$ Pg C yr$^{-1}$ throughout the 1980s and 1990s with little discernible trend. In their North America region they include both boreal and temperate land as well as some minor parts of tropical America. We find a sink of $-1.33$ and $-1.39$ Pg C yr$^{-1}$ for the 1980s and 1990s, respectively, for aggregated boreal and temperate North America. The implied trend is very small and agrees well with the fairly constant estimate of Potter et al. (2003). The considerable offset between decadal means is typical for comparisons of top-down and bottom-up methods (Gurney et al., 2004; Gurney and Eckels, 2011; Jacobson et al., 2007b). It is usually relieved by the incorporation of land priors, e.g. GE estimate a North American sink of $-0.97$ and $-0.59$ Pg C yr$^{-1}$ for the 1980s and 1990s, respectively, thereby halving the offset. However, they find a significant weakening of the sink in the 1990s that is not consistent with Potter et al. (2003).
Temperate north America accounted for -1.9 Pg C yr\(^{-1}\) of the 1980-2008 mean northern land sink, while Europe accounted for -0.4 Pg C yr\(^{-1}\). For 2000-08 we find a European sink of -0.33 (±0.64) Pg C yr\(^{-1}\) in good agreement with the CarbonTracker (Peters et al., 2007) estimate of -0.17 (±0.47) Pg C yr\(^{-1}\) averaged over 2001-07. Our 2000-08 sink for temperate north America of -2.17 (±1.17) Pg C yr\(^{-1}\) is far off the CarbonTracker estimate for 2001-07 of -0.62 (±0.42) Pg C yr\(^{-1}\), though still overlapping. The North American carbon budget from the SOCCR (Pacala et al., 2007) estimates a sink of -0.5 Pg C yr\(^{-1}\), in agreement with CarbonTracker but even further away from our estimate. When looking at our results for the different inversion modes (table 4.3), one can see that a large temperate north American sink is obtained by the unconstrained inversion in all decades. This sink then reduces in mode 2 due to the inclusion of the ocean prior. A further slight reduction is seen for mode 3 when averaging over the Stephens models. This pattern is persistent independent of the network used. We therefore conclude that a temperate north American sink as small as estimated by bottom-up studies can only be obtained when using land prior information that drive the sink down. For the 2000-08 period the relative contribution of ocean and land priors to reducing the sink can be demonstrated when comparing to the GE estimate:

1. unconstrained (mode 1): -3.19 Pg C yr\(^{-1}\)

2. control (mode 3, ocean prior applied): -2.17 Pg C yr\(^{-1}\)

3. GE (land prior applied): -0.86 Pg C yr\(^{-1}\)

4. SOCCR (bottom-up): -0.5 Pg C yr\(^{-1}\)

### 4.4.5 Tropical land

Decadal tropical source intensification is driven primarily by enhanced outgassing from Tropical America, whereas the flux from aggregated tropical Africa and Asia stays fairly constant - a pattern consistently found here and in GE's study.

One distinction is that GE find a slight decrease in tropical land source from the 1980s to the 1990s followed by a considerable increase, whereas we find a strong increase from the 1980s to the 1990s followed by a slight decrease in the 2000s. This pattern is consistent over all inversion modes and network choices, except for the 2000s decrease that can turn into no change or a small increase for certain setups (figure 4.9). It appears already in the unconstrained inversion, though accompanied with large uncertainties, hence it is a direct
4.4. Results and Discussion

consequence of atmospheric CO\textsubscript{2} evolution. Inclusion of the ocean prior dampens the interdecadal variations but leaves the general pattern untouched. Hence, the finding of GE of "mirrored" interdecadal changes follows from their use of land priors, though probably supported by the fact that tropical land flux estimates are generally the least certain (due to the lack of observations there) and can be adjusted by the inversion most easily to compensate for flux changes somewhere else.

While aggregated tropical Africa and Asia show no significant trend (neither in the GSNF, nor in the DSNF), there is a roughly constant offset of 0.8 Pg C yr\textsuperscript{-1} in decadal mean flux between the two studies: we find these regions to take up -0.41 Pg C yr\textsuperscript{-1} averaged over 1980-2008, whereas GE find a net release of 0.39 Pg C yr\textsuperscript{-1}. Whether tropical African and Asian land is characterized as net source or sink of carbon depends sensitively on the selection of transport models, as we identify a source, too, when averaging over all models. At the same time GE find a small sink when restricting to the three Stephens models. This model sensitivity was also apparent in the 1992-96 inversion in chapter 3, cf. discussion there and also in Gurney and Eckels (2011). In their recent synthesis Sarmiento et al. (2010) find the aggregated tropical African and Asian carbon budget to be balanced, with deforestation rate and vegetational sink cancelling each other. Their suggested net flux for the 1990s is 0.0 (±0.85) Pg C yr\textsuperscript{-1} (based on plot data from Chave et al. (2008) and using the extrapolation scheme of Phillips et al. (1998)). Though this overlaps with both our sink and GE’s source estimates, it leaves the question open whether tropical African and Asian regions act to release or take up carbon on an annual net basis (including emissions from deforestation and fires).

When looking at the regional level, it turns out that the absence of any trend in tropical uptake from Africa and Asia results from compensating trends in northern Africa and tropical Asia. The northern African region was nearly in balance in the 1980s before it turned into a sink in the following two decades. Its sink strength is estimated to be -0.50 Pg C yr\textsuperscript{-1} over both the 1990-99 and 2000-08 periods. The implied negative trend is partially offset by a positive trend in tropical Asia, which switched from a small sink in the 1980s to a nearly neutral balance in the following decades. The reason for the diagnosed strong model sensitivity of the aggregated tropical African and Asian trend lies in Northern Africa, where the trend becomes positive when averaged over all models (mode 2), contrasting to a robust trend for tropical Asia. Both regions exhibit strong changes in the 1980s mean flux when switching to the individual network setup. This, again, is related to the very sparse network in the 1980s that does not yet include CO\textsubscript{2} observations from a regular cruise in the western tropical Pacific (figure 4.2), which help to constrain fluxes from the adjacent land. The lack of these data is also expressed as large increases in uncertainty for the tropical Asian land flux estimate (table 4.4).

Tropical America is the dominant driver of the strong pantropical land source as well as its
intensification over the three decades. The region is estimated to act as a strong CO$_2$ source to the atmosphere throughout the examined period, increasing at a rate of 0.055 Pg C yr$^{-2}$. The decadal change between the 1980s and 1990s is more than 1 Pg C yr$^{-1}$, but in the 2000s the source actually weakens by 0.2 Pg C yr$^{-1}$.

In their study, GE also identified tropical America as the main driver of an increasing tropical land flux, though exhibiting a trend that is less than half of our estimate. Their decadal mean flux estimates are shifted towards a much weaker source or even a small sink, depending on which transport models are selected. The average offset over 1980-2008 is 3.0 Pg C yr$^{-1}$. Their trend curve is much more linear than ours, in particular they do not see the negative trend in the 2000s.

Trend estimates for tropical America are always positive, irrespective of whether or not we included the ocean prior, whether or not we averaged over the Stephens subset of models, and which network we used. However, interdecadal flux changes vary among the different setups. Using individual networks generally strengthens the linear trend but does not alter the pattern of interdecadal changes, i.e. whether the source increases or decreases between each two decades. Our finding of a decreasing flux in the 2000s is caused solely by the ocean prior, as can be seen by comparing the unconstrained results with those obtained with ocean prior included (figure 4.9). One potential explanation could be a reduction of deforestation-related fire emissions in the Amazonian region in 2008, as was reported by Le Quere et al. (2009) based on data of PRODES (2009). The 2008 emissions were 0.3 Pg C yr$^{-1}$ less than average for the preceding ten years. Our tropical American source reduction of 0.2 Pg C yr$^{-1}$ is consistent with that. However, considered that our estimate actually reflects the 2000-08 average and that the attached uncertainties are generally large in the tropical American region, the question whether deforestation emission cuts are driving our observed source reduction needs further investigation. With their use of land priors GE do not find a weakening in the tropical American source in the 2000s, but a considerable intensification.

Between the 1980s and 1990s book-keeping methods find only little increase in land-use change emissions in tropical America (e.g. Houghton (2003)). In a recent study based on the ecosystem model ORCHIDEE, Piao et al. (2009) estimate an increase in land-use change emissions of 0.1 Pg C yr$^{-1}$. With additional incorporation of rising atmospheric CO$_2$ and climate change they estimate a shift of 0.24 Pg C yr$^{-1}$ in the tropical American carbon exchange, consistent with the positive trend found here and by GE between the 1980s and the 1990s. However, we estimate a much stronger flux increase between these two decades, which cannot be explained by the above processes alone. To pin down what actually causes the large increase during this period we take a look at the DSNF and GSNF trends in tropical America: there is little change in the GSNF but a large increase in the DSNF of 1.46 Pg C yr$^{-1}$ in the 1990s that drives nearly all of the estimated net source intensification (figure 4.10b). As it is defined here by positive fluxes, the dormant season corresponds roughly to the wet
4.4. Results and Discussion

season of the southern parts of the Amazon rain forest. Unfortunately, owing to the Transcom region definitions the wetter northwestern Amazon cannot be separated from the more seasonal and drier southeastern Amazon. Nevertheless, the DSNF time period corresponds well to the Amazon-mean wet season due to the southern dominance of the seasonal signal. Our finding of a DSNF increase thus implies a shift in the balance of ecosystem respiration and gross primary production towards elevated carbon release during the Amazon-mean wet season. With our top-down approach we are not able to assess whether the DSNF increase is due to an increase in respiration, a decrease in productivity, or both. Based on four years of eddy covariance flux data in an evergreen, old-growth Amazonian rain forest, Hutyra et al. (2007) also find net carbon loss to the atmosphere during the wet season. They find increases in both productivity and respiration during the wet season, with the increase in respiration being dominant and driving the net carbon release.

There is emerging evidence that the Amazonian rain forest by itself plays a critical role in setting the onset of the wet season, thereby overlaying the natural oscillation of the tropical rain belt (e.g. Myneni et al. (2007)). An increase in evapotranspiration is observed towards the end of the dry season, which appears to increase the buoyancy of surface air and subsequently to increase the probability of atmospheric convection and rainfall. Myneni et al. (2007) find a 25% increase in LAI (Leaf Area Index) during the dry season over large parts of the Amazon rain forest. This is likely the cause for the enhanced evapotranspiration and supports the hypothesis that vegetation plays an important role in initiating the wet season. Although it is likely that such a seasonal signal in leaf area would lead to changes in canopy photosynthesis and ecosystem respiration, it is not clear whether it would lead to an increase or decrease in the net carbon exchange with the atmosphere, because of the complex interplay with other environmental and ecological constraints like vapor pressure deficits, temperatures, water and nutrient availability. Here we find a clear signal of elevated CO$_2$ release in August-September (end of dry season) for the 1990s and 2000s decades. This suggests a much stronger effect on respiration than on photosynthetic productivity.

4.4.6 Southern land

As mentioned before, the southern land trend over the past three decades is the strongest among the three latitudinal bands. However, after switching from near-neutral carbon exchange in the 1980s to a strong sink of -1.67 (±0.70) Pg C yr$^{-1}$ in the 1990s, the sink strength reduced in the 2000s to -1.20 (±0.93) Pg C yr$^{-1}$. These interdecadal changes do not take place homogeneously among the southern land regions, as will be discussed in the following. Because the contribution from Australia to southern land decadal fluxes and trends is generally small compared to those from southern America and Africa (figure 4.8 and table 4.2), we will not discuss that region in detail here.
Our results indicate that changes in the carbon budget in southern land regions are mainly responsible for the elevated land uptake in the 1990s compared to the 1980s. Or in other words, most of the additional CO$_2$ taken up by the land biosphere was actually taken up in the southern hemisphere, and only small parts by the northern hemisphere boreal and temperate regions. This agrees to the findings of Raupach (2011), which are based on the study of GE, who also identified only a weak northern land signal.

GE found the south African region to be the main driver for southern land decadal trends, which is also in remarkable agreement with our results, given the fact they used land priors to constrain their inversion, whereas we left the land completely free. The south African trend may be explained by higher rainfall in the region, leading to increased primary productivity. However, this should be visible in the seasonal cycle as an increase in growing season uptake (GSNF). While we find such an increase, it is only very weak and not significant, due to high monthly uncertainty and the generally rather aseasonal flux cycle. The latter results from the fact that the south African region is actually covered to a substantial degree by tropical forests, as it extends northwards to the equator (as opposed to south temperate America and Australia, cf. regional boundaries in figure 4.2).

For the 1990s, about 2/3 of the large intensification of the southern land sink is due to southern Africa and 1/3 due to south (temperate) America. The southern African region switched from a source to near-neutral carbon exchange, driven by a combination of enhanced growing season uptake (GSNF trend -0.065 Pg C yr$^{-2}$) and a decline in dormant season out-gassing (DSNF trend -0.075 Pg C yr$^{-2}$). The south American uptake increased mainly due to an intensification of the GSNF caused by higher peak uptake during the growing season as well as an earlier onset of the growing season (October-November instead of November-December).

By contrast, the weakening of southern land uptake in the 2000s is driven solely by the southern African region, which switches back to a net source. Again the GSNF and DSNF trends have the same sign and are comparable in size, causing the southern African seasonal flux curve to shift as a whole. While this could be driven by concurrent changes in photosynthetic productivity and respiration due to changes in temperature and rainfall patterns, it is difficult to interpret in view of the large uncertainties associated with African flux estimates. As we did not apply monthly land priors, the southern African seasonal cycle is poorly constrained, because the ocean prior’s ability to shape it is limited. The ocean prior constrains decadal means better through its well-defined flux estimates from adjacent ocean regions (tropical Indian and Atlantic, Southern Ocean). The shift of the southern African seasonal cycle as a whole could therefore simply reflect the inversion’s inability to constrain the region individually, and to adjust it in response to the adjacent longterm mean ocean fluxes.

Given the large uncertainties associated with southern Africa, the decadal trend estimates
4.4. Results and Discussion

are remarkably robust against network choice and the selection of transport models (mode 2 vs. mode 3, see figure 4.9). However, when only atmospheric CO$_2$ is incorporated (mode 1), the decadal trends show the opposite behavior: a slight increase from the 1980s to the 1990s followed by a decrease in the 2000s. Of course decadal mean fluxes are not significantly different from the other modes due to the very large uncertainties associated with the unconstrained inversion. The interdecadal pattern is clearly shaped by the ocean prior (inverted compared to mode 1) with little influence from the transport model selection. Interestingly, the indirect ocean constraints via the adjacent basins impose an interdecadal pattern and decadal trend very similar to GE’s estimates. As GE suffered from the same data scarcity over Africa as we did, it seems reasonable to assume that their southern African flux is predominantly defined by their land prior. From this it follows that ocean and land priors are consistent in this region in the sense that they constrain flux trends conformably.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>individual</td>
<td>common</td>
<td>individual</td>
<td>common</td>
<td>individual</td>
<td>common</td>
<td>individual</td>
<td>common</td>
</tr>
<tr>
<td>Bor N America</td>
<td>0.49</td>
<td>0.56</td>
<td>-0.10</td>
<td>0.54</td>
<td>0.41</td>
<td>0.37</td>
<td>-0.10</td>
<td>0.54</td>
</tr>
<tr>
<td>Boreal Asia</td>
<td>-0.91</td>
<td>-0.12</td>
<td>-0.05</td>
<td>0.42</td>
<td>-0.16</td>
<td>-0.50</td>
<td>-0.36</td>
<td>-0.26</td>
</tr>
<tr>
<td>Temp N America</td>
<td>-1.80</td>
<td>-1.37</td>
<td>-1.20</td>
<td>-1.88</td>
<td>-1.12</td>
<td>-1.92</td>
<td>-1.68</td>
<td>-1.73</td>
</tr>
<tr>
<td>Europe</td>
<td>-2.01</td>
<td>-1.22</td>
<td>-0.86</td>
<td>-1.26</td>
<td>-0.59</td>
<td>-0.10</td>
<td>-0.50</td>
<td>-0.10</td>
</tr>
<tr>
<td>Temp Asia</td>
<td>5.76</td>
<td>-0.57</td>
<td>-0.13</td>
<td>0.09</td>
<td>-0.53</td>
<td>0.00</td>
<td>-0.84</td>
<td>-0.49</td>
</tr>
<tr>
<td>Trop Africa</td>
<td>0.34</td>
<td>-0.37</td>
<td>0.65</td>
<td>0.70</td>
<td>0.68</td>
<td>2.25</td>
<td>-0.18</td>
<td>-0.50</td>
</tr>
<tr>
<td>Northern Africa</td>
<td>9.62</td>
<td>0.23</td>
<td>-1.06</td>
<td>1.06</td>
<td>0.78</td>
<td>-0.09</td>
<td>0.72</td>
<td>0.55</td>
</tr>
<tr>
<td>Trop Asia</td>
<td>-10.39</td>
<td>0.68</td>
<td>0.66</td>
<td>-0.47</td>
<td>-0.65</td>
<td>0.51</td>
<td>0.24</td>
<td>0.07</td>
</tr>
<tr>
<td>South America</td>
<td>-16.32</td>
<td>1.38</td>
<td>-0.09</td>
<td>-0.87</td>
<td>0.07</td>
<td>-1.21</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>10.66</td>
<td>0.86</td>
<td>2.01</td>
<td>1.44</td>
<td>0.57</td>
<td>0.38</td>
<td>1.66</td>
<td>0.52</td>
</tr>
<tr>
<td>Australasia</td>
<td>-0.58</td>
<td>-0.52</td>
<td>-0.48</td>
<td>-0.23</td>
<td>-0.21</td>
<td>-0.17</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>Northern Ocean</td>
<td>0.30</td>
<td>-0.18</td>
<td>-0.18</td>
<td>-0.05</td>
<td>-0.18</td>
<td>0.00</td>
<td>-0.16</td>
<td>0.18</td>
</tr>
<tr>
<td>N Pacific</td>
<td>-0.78</td>
<td>-0.43</td>
<td>-0.43</td>
<td>-0.68</td>
<td>-0.43</td>
<td>0.00</td>
<td>-0.40</td>
<td>0.43</td>
</tr>
<tr>
<td>N Atlantic</td>
<td>-0.32</td>
<td>-0.33</td>
<td>-0.33</td>
<td>0.36</td>
<td>-0.33</td>
<td>0.00</td>
<td>0.41</td>
<td>-0.34</td>
</tr>
<tr>
<td>W Pacific</td>
<td>2.82</td>
<td>0.09</td>
<td>0.09</td>
<td>0.13</td>
<td>0.09</td>
<td>0.00</td>
<td>-0.34</td>
<td>0.09</td>
</tr>
<tr>
<td>E Pacific</td>
<td>1.90</td>
<td>0.24</td>
<td>0.24</td>
<td>0.55</td>
<td>0.24</td>
<td>0.00</td>
<td>0.53</td>
<td>0.23</td>
</tr>
<tr>
<td>Trop Atlantic</td>
<td>-8.26</td>
<td>0.16</td>
<td>0.16</td>
<td>-0.82</td>
<td>0.16</td>
<td>0.00</td>
<td>-1.65</td>
<td>0.15</td>
</tr>
<tr>
<td>Trop Indian</td>
<td>-2.00</td>
<td>0.10</td>
<td>0.10</td>
<td>0.43</td>
<td>0.10</td>
<td>0.00</td>
<td>-0.75</td>
<td>0.10</td>
</tr>
<tr>
<td>S Pacific</td>
<td>0.98</td>
<td>-0.45</td>
<td>-0.45</td>
<td>-0.45</td>
<td>-0.45</td>
<td>0.00</td>
<td>-1.42</td>
<td>-0.46</td>
</tr>
<tr>
<td>S Atlantic</td>
<td>7.38</td>
<td>-0.19</td>
<td>-0.19</td>
<td>0.76</td>
<td>-0.19</td>
<td>0.00</td>
<td>-0.32</td>
<td>-0.19</td>
</tr>
<tr>
<td>S Indian</td>
<td>0.33</td>
<td>-0.48</td>
<td>-0.48</td>
<td>-0.68</td>
<td>-0.48</td>
<td>0.00</td>
<td>-0.88</td>
<td>-0.49</td>
</tr>
<tr>
<td>Southern Ocean</td>
<td>0.56</td>
<td>-0.27</td>
<td>-0.27</td>
<td>-0.28</td>
<td>-0.28</td>
<td>0.00</td>
<td>-0.28</td>
<td>-0.28</td>
</tr>
<tr>
<td>Northern Land</td>
<td>1.53</td>
<td>-2.73</td>
<td>-2.14</td>
<td>-3.14</td>
<td>-2.75</td>
<td>-2.41</td>
<td>-2.10</td>
<td>-2.83</td>
</tr>
<tr>
<td>Tropical Land</td>
<td>-0.43</td>
<td>0.54</td>
<td>0.25</td>
<td>1.29</td>
<td>1.81</td>
<td>1.92</td>
<td>0.34</td>
<td>2.53</td>
</tr>
<tr>
<td>Trop Africa+Asia</td>
<td>-0.77</td>
<td>0.91</td>
<td>-0.39</td>
<td>0.59</td>
<td>0.13</td>
<td>-0.34</td>
<td>0.53</td>
<td>0.30</td>
</tr>
<tr>
<td>Southern Land</td>
<td>-6.23</td>
<td>1.71</td>
<td>1.43</td>
<td>0.34</td>
<td>0.43</td>
<td>0.00</td>
<td>0.71</td>
<td>1.16</td>
</tr>
<tr>
<td>Trop+South Land</td>
<td>-6.67</td>
<td>2.25</td>
<td>1.68</td>
<td>1.62</td>
<td>2.24</td>
<td>1.92</td>
<td>1.05</td>
<td>1.83</td>
</tr>
<tr>
<td>Northern Ocean</td>
<td>-0.80</td>
<td>-0.94</td>
<td>-0.94</td>
<td>-0.37</td>
<td>-0.94</td>
<td>0.00</td>
<td>-0.54</td>
<td>-0.95</td>
</tr>
<tr>
<td>Tropical Ocean</td>
<td>-5.53</td>
<td>0.59</td>
<td>0.59</td>
<td>0.29</td>
<td>0.59</td>
<td>0.59</td>
<td>0.00</td>
<td>0.57</td>
</tr>
<tr>
<td>Ocean in South</td>
<td>9.26</td>
<td>-1.39</td>
<td>-0.64</td>
<td>-1.39</td>
<td>-1.39</td>
<td>0.00</td>
<td>-2.21</td>
<td>-1.42</td>
</tr>
<tr>
<td>Global Land</td>
<td>-5.13</td>
<td>-0.48</td>
<td>-0.46</td>
<td>-1.51</td>
<td>-0.51</td>
<td>-0.49</td>
<td>-1.10</td>
<td>-0.49</td>
</tr>
<tr>
<td>Global Ocean</td>
<td>2.92</td>
<td>-1.74</td>
<td>-1.74</td>
<td>-0.73</td>
<td>-1.74</td>
<td>-1.74</td>
<td>0.00</td>
<td>-2.15</td>
</tr>
<tr>
<td>Global Total</td>
<td>-2.21</td>
<td>-2.22</td>
<td>-2.22</td>
<td>-2.24</td>
<td>-2.24</td>
<td>-2.23</td>
<td>0.00</td>
<td>-3.08</td>
</tr>
</tbody>
</table>

Table 4.3 Annual posterior fluxes for all 22 regions in the inversion as well as selected aggregated regions. Control results (mode 3, common network) are compared to the study of Gurney and Eckels (2011) (GE), as well as to the results from all modes and both networks. Bold numbers indicate 90% statistical significance for the flux sign. Explicit uncertainties are omitted for clarity here, but can be found in table 4.4.
### Table 4.4: Annual posterior total uncertainties for all fluxes in the inversion as well as selected aggregated regions.

Flux values are tabulated in table 4.3.
4.5 Summary and conclusions

In this paper we present seasonal fluxes and decadal trends of regional carbon exchange averaged over three periods: the 1980s and 1990s decades and 2000-08. Monthly mean fluxes are obtained through a multi-step Bayesian inversion in three modes. Mode 1 is driven by atmospheric CO$_2$ records alone, whereas modes 2 and 3 include additional information on oceanic exchange flux and covariance, derived from a separate ocean interior DIC inversion combined with surface ocean pCO$_2$ data. In mode 3 we selected three out of eleven atmospheric transport models according to their skill in simulating vertical gradients of CO$_2$. For each mode and decade we run the inversion on two different observational networks, a common network containing the same stations for each decade and an individual network, where stations were selected according to record density over the decade. The common network contains some stations with insufficient data coverage during the 1980s. We assigned large uncertainties to these stations to keep their influence small, however, they impose some regularization on flux estimates. The setup with common stations and all constraints represents our control inversion.

Oceanic trends are well-constrained here, because the underlying ocean inversion derives fluxes with small uncertainties as a consequence of the vast amount of DIC data available. We did not consider interannual variations in oceanic transport, so that oceanic trends follow directly from the anthropogenic perturbation in atmospheric CO$_2$. This could introduce an additional error source, as concluded by Roedenbeck et al. (2003a). While the effect is likely to be small for the global ocean trend (-0.007 Pg C yr$^{-2}$), there are indications of climate change induced uptake saturation in some regions, most prominently in the Southern Ocean (Le Quere et al., 2007; Lovenduski et al., 2008). In an experiment we have shown that fluxes from some land regions respond sensitively to an assumed cessation of the Southern Ocean sink, in particular in southern Africa and America.

The global trend of combined oceanic and land uptake is estimated to -0.055 Pg C yr$^{-2}$ (increasing uptake), based on atmospheric CO$_2$ growth rate and fossil fuel emission rates. Global land areas contribute 87% to this trend and the global oceans 13%. The dominant role of the land for the global trend is also apparent when looking at decadal means: in the 1980s the contribution from the land biosphere to the global sink was 22% before it doubled to 43% in the 2000s. Most of this occurred already during the 1990s, in agreement with bottom-up estimates of Le Quere et al. (2009) and Sarmiento et al. (2010), who attribute the rapid increase in part to the Pinatubo eruption in 1991. In their recent interannual atmosphere inversion Gurney and Eckels (2011) do not see such a signal, but a much more linear increase in land uptake. Because they used an inversion setup similar to ours, though with land prior and only very weak ocean prior, we conclude that it is our ocean prior that makes our land sink consistent with bottom-up estimates. The land sink is consistent also
with regard to absolute magnitude, as can be seen by comparison to the recent ecosystem model study of Piao et al. (2009).

Despite the lack of interannually varying ocean circulation we deem our ocean prior more realistic than leaving the ocean unconstrained (as done by Gurney and Eckels (2011)), because the atmospheric inversion otherwise tends to misallocate land fluxes to adjacent oceanic regions. The latter leads to unrealistic oceanic trends (for example a positive trend, i.e. weakening uptake) and may subsequently cause unrealistic land trends.

Decadal trends and variations are generally less sensitive to changes in network composition and inversion mode than decadal mean fluxes, in agreement with previous findings of Gurney et al. (2008) and Roedenbeck et al. (2003a). Exceptions include the African continent where the inclusion of the ocean prior plays the dominant role, and South America where more pronounced trends are found with the individual network. Fluxes from Africa are poorly constrained by atmospheric observations (scarce data) and, hence, are mainly determined indirectly via adjacent oceanic regions, which are well constrained by the ocean prior. For South America the switch between networks leads to a redistribution of flux between the tropical and temperate regions, in particular in the 1980s.

The ocean prior plays a key role in setting trends in all latitudinally aggregated land areas, i.e. northern, tropical and southern land. The impact of network and model choice is less significant. While the uptake from northern land has increased over the decades at a rate of 0.034 Pg C yr$^{-2}$, its relative contribution to the global land sink has decreased by about 20%. By contrast, the contribution from southern land increased notably. The estimated rapid increase in land uptake between the 1980s and 1990s is in large parts due to southern land regions. Southern Africa is the main driver of the enhanced southern land uptake between the 1980s and 1990s, which may be caused by more intense precipitation in this region and consequently higher NPP.

Northern extratropical land regions removed about 2.7 Pg C yr$^{-1}$ from the atmosphere averaged over the whole 1980-2008 period. This estimate is somewhat larger than those of Gurney and Eckels (2011) (2.0 Pg C yr$^{-1}$) and Ciais et al. (2010) (1.7 Pg C yr$^{-1}$ for 2000-04), though still overlapping within the uncertainty bounds. The high rate is driven by strong uptake from temperate north America, accounting for 70% of the total. By contrast, bottom-up studies such as the SOCCR (Pacala et al., 2007) suggest much smaller uptake. We showed that the temperate north American sink shifts more and more towards bottom-up estimates as we include more constraints. However, given the underlying CO$_2$ observations and Transcom models it is only possible to converge further if explicit land prior information is incorporated.

At the regional level we find weakening European and intensifying temperate north American sinks, therefore exchanging the roles of both sub-continents compared to the land prior-based atmospheric inversion study of Gurney and Eckels (2011). While the sign of both
trends is robust against changes in network and transport, its strength responds sensitively
to network changes.

We find a negative trend in boreal land flux that explains 79% of the overall northern land
trend. While trends are of equal magnitude in the boreal parts of America and Asia, their sea-
sonal signatures are different. A more intense growing season net uptake drives the boreal
American trend, likely caused by higher photosynthetic activity due to warmer temperatures.
Boreal Asia is characterized by a decline in seasonal amplitude, i.e. a weakening in both
dormant and growing season net fluxes.

Tropical land is estimated to release large amounts of carbon, about 2.6 Pg C yr$^{-1}$ on average
over the three decades (land-use change emissions included), at an increasing rate of 0.048
Pg C yr$^{-2}$. Both the source magnitude and trend are driven entirely by tropical America, while
the combined tropical regions in Africa and Asia act as net sink and exhibit almost no trend.

A strong source from tropical America is imposed indirectly by constraining the adjacent
oceanic regions according to our ocean prior. The source is further enhanced by the applied
model selection, because of ventilation characteristics of the three Stephens models. Both
findings were already discussed in chapter 3 and are observed here as well. Without regular-
ization by land priors the tropical American source will always exceed bottom-up estimates,
such as flux tower and inventory studies, which usually find the region nearly in balance or a
weak sink of carbon.

1980 to 2008 trend estimates for tropical America are always positive, irrespective of network
and inversion mode. However, we find a decrease in source of 0.2 Pg C yr$^{-1}$ between the
1990s and 2000s that is completely attributable to the ocean constraint. Such a decrease is in
agreement with a sudden reduction of deforestation-induced fire emissions in the Amazonian
region in 2008 (Le Quere et al., 2009), though large uncertainties and the fact we only see
the 2000-08 mean make a clear attribution difficult.

Our finding of a large increase in the tropical American source between the 1980s and 1990s
is driven by enhanced outgassing during the Amazon-mean wet season. Our top-down ap-
proach does not allow to attribute this to increased respiration or decreased photosynthetic
productivity, though Hutyra et al. (2007) suggest increases in both productivity and respi-
ration with the latter being dominant. We identify an abrupt transition towards outgassing
at the end of the dry season, coinciding with an increase in leaf area observed by Myneni
et al. (2007) that was proposed to play an important role in the initiation of the wet season.
Our results, hence, suggest that the enhanced leaf growth does not only lead to increased
evapotranspiration but also to elevated CO$_2$ release.

Southern extratropical land regions play a key role for the overall land uptake trend. Their sink
strength increased at a rate of -0.062 Pg C yr$^{-2}$. Especially between the 1980s and 1990s
most of the additional CO$_2$ taken up by the land biosphere was actually taken up by southern land regions. Northern boreal and temperate regions contributed only a small fraction, as was concluded by Raupach (2011) as well.

The main driver for southern land decadal variability is southern Africa. We find this region to have switched from a source in the 1980s to near-neutral carbon exchange in the 1990s and back to a (smaller) source in the 2000s. The reason could be changes in rainfall patterns, though we do not find significant seasonal signatures due to large monthly flux uncertainty. Despite the uncertainties, the decadal trend estimates are insensitive with respect to changes in network and inversion mode. The pattern of decadal variability is primarily shaped by the ocean constraint and is in good agreement with estimates of Gurney and Eckels (2011). We therefore conclude that ocean and land priors impose consistent constraints on decadal variability in this region.
Chapter 5

Summary and outlook

Over the period of my Ph.D. project, I have developed a Bayesian synthesis inversion to estimate spatially distributed CO$_2$ fluxes across the air-land and air-sea interfaces that are jointly constrained by carbon data from the atmosphere and the ocean. The novelty of my work can be summarized to the following key features:

- Multiple data streams from the atmosphere and the ocean are incorporated, in some cases augmented with constraints from the land. The impact of each constraint on flux estimates is quantified. In addition, the sensitivity to changes in atmospheric transport and observational network is assessed.

- The joint approach has seasonal resolution and is applied to various time periods between 1980 and 2008, thus enabling the estimation of decadal trends in the CO$_2$ fluxes.

I want to point out that none of these features would, by itself, be a completely new development. Rather, it is the simultaneous integration of all features, as well as the exploration of their coupling and interaction, that constitutes a unique contribution to the scientific work on the estimation of CO$_2$ fluxes.

For example, Jacobson et al. (2007a,b) developed a joint ocean-atmosphere inversion as well, but remained limited to annual mean fluxes. Furthermore, they did not address network sensitivity or decadal variability. Others resolved CO$_2$ fluxes on shorter temporal scales, such as monthly (Gurney et al., 2004) or even weekly (Peters et al., 2010, 2007, CarbonTracker), but either utilized atmospheric CO$_2$ as the only data stream or were limited to regional scales. Yet others assessed the sensitivity to modeled atmospheric transport (Stephens et al., 2007) or observational network composition (Gurney et al., 2008; Roedenbeck et al., 2003b), but excluded the constraints from ocean carbon data. Similarly, some studies addressed inter-annual variability in the CO$_2$ fluxes, but again based on atmospheric CO$_2$ data only. This can result in unrealistic air-sea flux estimates and trends, with the study of Gurney and Eckels (2011) as an example, because atmosphere-only inversions tend to partially outsource
the strong terrestrial seasonal flux signal to adjacent ocean regions. As a final example, Le Quere et al. (2009) produced CO$_2$ flux estimates with interannual resolution covering the last 50 years. However, their work is restricted to global total fluxes, and based on land and ocean process models without considering any direct data stream.

By contrast, flux estimates from my joint inversion are simultaneously consistent with (at least) three data streams: atmospheric CO$_2$, ocean interior DIC and surface ocean pCO$_2$.

### 5.1 Overview of main results

Naturally, the ocean data provide a strong constraint for air-sea CO$_2$ fluxes, yet their strength is much greater on annual than on monthly timescales, owing to the longterm character of the flux estimates from the ocean interior DIC inversion. On annual scales they also provide a considerably strong constraint for air-land CO$_2$ fluxes, because they support the atmosphere inversion in distinguishing between flux contributions from neighbored land and ocean regions. On monthly scales they do not play a major role in shaping the terrestrial seasonal cycle, but they prevent the atmosphere inversion from allocating unrealistic seasonal patterns to air-sea fluxes. I demonstrated the annual strength of the ocean prior also in the context of a pseudo-inversion of remotely sensed atmospheric column mean CO$_2$ data: if included in the pseudo-inversion, the minimal required accuracy of the column mean concentrations to catch up with the surface network increased to 1.2 ppm or better, compared to 2 ppm without the ocean prior. Currently active or proposed satellite missions dedicated to CO$_2$ measurements aim at accuracies between 1 and 1.5 ppm, so just at the edge of competing with the combined data stream approach.

The strength of each constraint in the joint inversion is considerable with regard to absolute fluxes in the northern, tropical and southern latitudinal bands: omitting one constraint leads to non-negligible changes in flux values and uncertainty. This does not hold, in general, for the estimation of decadal flux variability: decadal trends in aggregated northern and southern land regions are influenced most by the ocean prior and less by model selection and network composition. In the tropics the model and network influences become significant. For some regions, like the northern land, all constraints drive the results in the same direction. Yet there are regions where different constraints pull the results in opposite directions: the greatest effect is in the tropical American region, where the ocean prior and the Stephens models imply a strong source, while the inventory-based annual land constraint points towards a near-zero flux. The inconsistency of constraints for tropical America holds for all inversion periods. Without additional regularization by explicit land prior fluxes, tropical America will always be estimated to be a strong source of carbon, in disagreement with most bottom-up estimates.
The strong tropical American source is the reason for a pantropical land source. Yet, regarding the pantropics the Stephens models estimate a smaller source than the all-model mean. This implies a weaker CO$_2$ uptake by northern land ecosystems due to mass balance requirements communicated through atmospheric transport. Looking at the seasonal cycle reveals: the weaker northern land uptake is driven by stronger winter respiration that dominates over a concurrent increase in summer uptake (implying an enhanced seasonal amplitude).

Between the 1980s and the 1990s a remarkable jump in global land uptake occurred, increasing from -0.5 to -1.2 Pg C yr$^{-1}$. This jump was also clearly identified by Sarmiento et al. (2010), who attribute it in part to the Pinatubo eruption in 1991, but also point out that it likely started already in the late 1980s. Le Quere et al. (2009) also find this rapid sink increase with their bottom-up approach. My joint inversion largely attributes the jump to the southern land regions, in particular to southern Africa. Interestingly, Gurney and Eckels (2011) do not find this jump, but rather a more linear increase over the 1980-2008 period. In their interannual inversion, the post-Pinatubo increase in land uptake is offset by a reduced uptake in the second half of the 1990s (partially caused by the 1997/8 El Nino). Since the rapid increase in global land uptake is implied by the well-constrained ocean sink trend, I conclude that the ocean prior agrees very well with bottom-up land estimates at the global scale.

Boreal regions dominate the decadal flux trend over northern land. The increase in boreal land uptake between 1980 and 2008 is equally distributed among the boreal American and Asian regions, yet the seasonal drivers are different: boreal America shows a more intense growing season uptake (possibly due to warmer temperatures and thus enhanced photosynthesis), while boreal Asia shows a weakening of growing season uptake accompanied by less dormant season outgassing.

5.2 Caveats and outlook

My approach is a monthly cyclostationary inversion, that is, monthly fluxes are estimated, yet they represent the average over 5 (for the 1992-1996 inversion) or 10 years (for the decadal inversions; 9 years for the 2000-2008 inversion). Interannual variability is not considered. An interannual inversion would have been possible in principle with the Transcom suite of atmospheric transport models, as has been done by Baker et al. (2006); Gurney et al. (2008); Gurney and Eckels (2011). But an interannual joint ocean-atmosphere inversion was not possible, because of the limited temporal resolution of the ocean constraint.

Air-sea flux estimates from the ocean inversion represent longterm (several decades) means only. I scaled them to represent annual mean fluxes for each desired year, but this scaling
is based on the atmospheric CO$_2$ perturbation alone, in particular no interannually varying ocean circulation was included. Rather, the 10 ocean circulation models used in the ocean inversion were driven by re-analyzed winds and surface forcing. As a result, inverse ocean fluxes do not resolve interannual variations, such as induced by ENSO or climate change. Ocean models with interannually varying circulation are available, and it has been shown that the effect on air-sea CO$_2$ fluxes can be noticeable (Le Quere et al., 2009; Lovenduski et al., 2008). For example, there are indications for a saturation of the Southern Ocean sink. For the global ocean sink, the models reviewed by Le Quere et al. (2009) show interannual variations that are more pronounced than those based on my anthropogenic scaling, yet they are still very small compared to terrestrial interannual variability. Furthermore, when averaged over a decade the differences are marginal. Nevertheless, some models also show the beginning of a saturation of the global ocean sink over the last 2 to 3 decades, possibly due to climate change, which the ocean inversion cannot detect. So why did I use ocean models with stationary circulation instead of interannual ones? The main reason is because there is no suite of interannual models of which response fields for an interior dye tracer are available. There is a trade-off between using a single interannual ocean model for the inversion, and using multiple stationary models. The multi-model approach has the advantage that the transport error due to imperfect circulation and forcings can be assessed. This assessment may still be incomplete, because all models may exhibit similar structural deficiencies, but with a single model approach no such assessment would be possible.

I used surface ocean pCO$_2$-based observations (Takahashi et al., 2009a) to seasonalize the scaled annual ocean inverse fluxes. The extraction of seasonal patterns was based on the pCO$_2$ climatology, not the raw measurements. The climatology deemed appropriate for cyclostationary inversions over 5 to 10 years, and it allowed for higher data density. Yet, Takahashi et al. referenced the climatological pCO$_2$ data to the year 2000 by applying a global and constant annual rate. The direct usage of raw measurements at their original coordinates in space and time would be preferable, in particular one could think of mapping an interannually resolved flux pattern on the ocean inverse result. The limiting factor in this regard is, however, the data density (in particular the temporal density). In my mapping scheme I based the estimation of monthly flux uncertainty on the pCO$_2$ data density; if I applied that scheme to interannually resolved data the resulting uncertainties would increase so much that the air-sea fluxes would remain virtually unconstrained in the joint inversion. In the future the development of other more direct techniques for the assimilation of pCO$_2$ data are certainly of interest. One such technique is currently being developed in the framework of the CARBONES$^1$ project (see figure 5.1 for an overview of the project).

Similar to the ocean models, the atmospheric models did not provide interannually varying transport. Modelers chose a particular calendar year of transport winds and recycled those

---

$^1$CARBONES project: 30-year re-analysis of CARBON fluxES and pools over Europe and the globe. See http://www.carbones.eu
5.2. Caveats and outlook

Figure 5.1 Schematic view of the CARBONES data assimilation system. The system comprises a land ecosystem model (ORCHIDEE), an ocean biogeochemical model (PISCES) and an atmospheric transport model (LMDZ). Parameters of all the models are optimized with regard to several observational constraints: atmospheric CO\textsubscript{2} from satellites and surface stations, direct CO\textsubscript{2} flux measurements from towers, and carbon inventories from forests and soils.

across the years of simulation of the response functions. Different modeling groups utilized different chosen years for their recycled transport (table 4.1), though it is not clear what impact this choice has on CO\textsubscript{2} flux results. Within the pool of atmospheric inverse studies with resolved interannual CO\textsubscript{2} fluxes, the influence of interannually varying transport remains inconclusive though the available evidence suggests that it is of secondary importance (Roe- denbeck et al., 2003a; Peylin et al., 2005; Piao et al., 2008). Clearly, it would be preferable to use atmospheric models with interannually varying transport for my joint inversion. But as for the ocean inversion, a trade-off exists between realistic interannual transport and the possibility to assess transport errors. The Transcom suite of more than 10 atmospheric models allows for such an error assessment. Future model intercomparison studies could, however, include interannually varying transport. The computation of the four-dimensional response fields for multiple models constitutes the most time-consuming part of any inversion and is, hence, the limiting factor on such project visions. Simply collecting model output from various modeling groups would not be sufficient, because the model runs need to be based on the same set of input fields. The need for an organized modeling protocol in concert with the computational challenges makes such an intercomparison project a large effort. Yet I think it would be a worthwhile effort. A first attempt has recently been made by Ciais et al. (2010) based on four different atmospheric inversions, however, the transport models underlying the inversions did not use common input fields.

My approach assumes the global CO\textsubscript{2} flux due to fossil fuel emissions to be a fixed known with no associated uncertainty. This means that any bias in the fossil fuel flux estimates and trends will be translated into the non-fossil fluxes and trends estimated. In addition, the spatial distribution of the fossil emissions is included for every year in the inversion periods, but is based on two maps for 1990 and 1995. After 1995 considerable changes occurred to the emission pattern, most importantly with regard to an increased fossil CO\textsubscript{2} source from China, which led to a bulge in the global zonal mean distribution (figure 1.5) closer to the equator. An accumulation of CO\textsubscript{2} in the PBL there would be misinterpreted by the inversion
as an additional near-tropical source from the terrestrial biosphere. I did not quantify as to which degree such an effect may have influenced my estimates for the tropical and temperate Asian fluxes.

A general limitation to atmospheric Bayesian synthesis inversions is that CO$_2$ flux attribution to ecosystem processes is difficult, because only net fluxes between the atmosphere and the Earth's surface can be estimated. Conclusions on processes can only be made indirectly, for example by analyzing the seasonal flux cycle. An alternative are integrated carbon cycle data assimilation systems, such as developed by CARBONES or CarbonTracker$^2$. They are similar to my approach in that they assimilate observational carbon data from multiple data streams, yet they do not optimize for net fluxes but for process-based model parameters. Given the optimized set of model parameters, it is possible to attribute flux results to the underlying modeled processes in the ocean or the land biosphere. Furthermore, it is possible to use these optimized models to make projections into the future. These projections include net CO$_2$ fluxes, but also all the information on changes in the modeled mechanisms. This is of particular interest not only from the scientific point of view, but also from a political and societal point of view, since it helps us move towards the goal of global monitoring of CO$_2$ sinks and sources in near real time.

Perhaps the main disadvantage of such integrated systems is the potential of structural biases in the process models. For example, if an important process is missing or incorrectly parameterized in the terrestrial ecosystem model, the whole assimilation system would be biased in response to its attempts to explain the observations. An alternative would be to improve Bayesian inversion systems, as they are independent from the modelization of processes. If it was possible to perform atmosphere or joint ocean-atmosphere inversions with high spatiotemporal resolution and low flux uncertainty, they could also be used as global monitoring systems. The attribution to processes would become easier if, for example, daily resolution is possible. However, such inversions place very high demands on the data to be inverted: atmospheric CO$_2$ must be measured from a very dense and global surface station network with quasi-continuous sampling. Extending the current network to such a network appears unrealistic due to high costs and areas difficult to access. Alternatively, remotely sensed CO$_2$ column data could be assimilated, yet the associated issues (particularly measurement accuracy) need to be resolved first.

Acknowledgements

I would like to thank my supervisor Nicolas Gruber for making this project possible and for guiding me through the past four years. The attractive make-up of the Ph.D. position offer immediately caught my eye, making the decision easy to move to Switzerland and also to move away a bit from the field of fundamental physics into the environmental physics field. I will never forget the "trick" to invite me to ETH and bring me right to the Polyterrasse on a sunny day with a fantastic view of the Alps. It made the decision even easier.

Funding for my research was partially provided by the Swiss Federal State Secretariat for Education and Research (SER) and coordinated by the European Cooperation in Science and Technology (COST). Partial funding was also provided by the FP7 European Community Programme in the framework of the CARBONES project. I am happy to be part of this project, especially as it works in a highly collaborative and instructive environment.

During my first summer school at the IFM-GEOMAR in Kiel I listened to a talk of Andreas Oschlies concerning model-data fusion concepts. From that talk and our subsequent discussion I learned much more about my own approach than I had been able to understand during the preceding three months, thank you!

I am also grateful to Christian Rödenbeck and Sara Mikaloff-Fletcher for sharing with me their knowledge about atmosphere and ocean inverse techniques. In particular, many positive memories remain from my walks with Christian in Utrecht, showing once more that a certain distance between the conference center and the hotel can be of great advantage.

I also want to thank Alina Freing for making my stay in Kiel very worthwhile and for facilitating my participation in a cruise in the Baltic Sea. Although it was a one-day cruise only, I was lucky (or unlucky?) to experience one of the worst local weather conditions of the past decade, associated with very rough seas. Despite the fun we had on the unstable ship, this trip will always stay in my memory as the day where I realized that each data point in a typical oceanographer’s excel sheet may represent a lot of work and suffering.

It has been a privilege to work in the young environmental physics group at ETH, to see the group growing from just 4 to more than 15 people. There were plenty of opportunities to work on myself and to grow in personality, thanks to the healthy mix of interesting people in the
group. I am grateful for the great time I had here and for the many ways you guys enriched my life! This is not limited to the office, since many of you became much more than colleagues to me. Olivier and Ilaria, thank you for giving me company during the final rush of writing this thesis! Although there are now a couple of restaurants I will probably never set foot in again, I sometimes even enjoyed the crazy life we lived during that period of time. Claudine! Thank you for being my officemate for so long! I especially enjoyed our constructions of complex and interwoven networks of rumor. I think you really deserve the two titles “Corsican dancing queen” and “Miss Alpine 2010”!

Finally, a big thank you to my friends and family who always stood behind me and endured my sometimes unsocial behavior towards the end of my Ph.D. project. Special thanks to Berni in this regard! I am really looking forward to the series of vacations we are going to undertake! I am particularly grateful to my parents for giving all of us - Lars, Tine, Anke and me - the possibility to freely choose our profession, for supporting our education all the way through, and to enable us to follow our own approach to life, despite the tight financial constraints. Vielen vielen Dank!!!
Appendix A

Cyclostationary CO$_2$ response fields

I applied some changes to the way how the cyclostationary model responses are computed compared to the original Transcom code. Their computation is done in the Matlab script "makefrech.m" (figures A.1, A.2 and A.3). Code snippets highlighted in red contain the changes as well as a brief explanation.

The key point is that the flux pulses from each region and month lasted only one month and were then discontinued for the remainder of the three years modeling period. By contrast, the presubtracted (background) fields were applied for one full year. Therefore, in order to get the cyclostationary response, it is correct to sum the annual responses to the background fields for each of the three years and then to detrend the responses by that fraction of the first year that corresponds to the time the background fields were applied. However, in case of the pulsed monthly fluxes, the detrending must consider three full years, because the pulse was already completed, independent of the month it was released.

For all my inversions I changed the detrending scheme accordingly.


```matlab
function frech = makefrech(invcon, outcon, modcon, data, ...  
   ndat, neqns, nsources, nreg, npreff, modnum)
% This script builds the equation part of the Frechet/Jacobi matrix for  
% the cyclostationary inversion. Additional a priori constraints are  
% added as well.
% All transport models contained in "modconfig" are considered.
% If a cyclostationary pseudo column data inversion is performed, this  
% routine calls the appropriate subroutine 'read_COLresponses'.
% First, the Green's functions (i.e. the model responses at the  
% observation sites) are built.
% From that the Frechet (i.e. the Jacobian) is built:  
% NOTE: each location in the code where the script could be extended in  
% order to support tc3tdi/interannual inversions is marked as:
% *** To be extended ***
% Author: Kay Steinkamp  
% Date: Nov 2008  
% Oct 2009: updated to include annual ocean OIP constraints  
% May 2010: updated to handle pseudo column data inversion  
% % get needed parameters  
% in read_con it was made sure that all necessary /fields exist  
% global monperyr monlen daysperyr presubGTC
nsites = data.nsites;  
npreff = invcon.npresub;oi_flag = (modnum == 1);  
% avoid double−dumping  
% % set and check parameters  
% ppmv = 0.47079;  
% conversion factor: 1 PgC = 0.47079 ppmv  
% elapm = int32(3*monperyr);  
% number of months written for the CO2 tracers  
% % check dimensions are correct  
% if npreff < npreff  
% error('makefrech: too many presubs in responses!')  
% else npreff = npreff+1  
% disp('npreff=preff+1, extra responce is offset.')  
% end  
% % compute mid−month time values (used for detrending)  
% tmidmon = zeros(monperyr,monperyr);  
% Note: each row of tmidmon corresponds to a starting month (pulse)  
% % (row 1 = jan, etc). From its starting month on, each row then  
% % contains mid−month values for any subsequent month within a year.  
% for i = 1:monperyr  
% kount = 0;  
% for c = 1:monperyr
```

This snippet of code seems to be part of a larger script that builds the equation part of the Frechet/Jacobi matrix for a cyclostationary inversion. It includes the building of Green's functions, which are essentially the model responses at the observation sites, and uses these to construct the Frechet matrix. The code also handles various constraints and checks the dimensions of the data and parameters to ensure that they are correct.
mpulse = monperyr; % clarity: month of pulse vs. month in year
fracyr = tmidmon./daysperyr; % mid-month expressed in year fraction
lenfrac = monlen./daysperyr; % month length relative to year length

% now the npresub responses (should be 4 in number)
% and the CO2 land/ocean responses (should be 12*22=264 in number)
ipre = 1:npresub; % presub indices
itrcp = ipre + monperyr*nreg; % assign presubs to tracers
ireg = 1:nreg; % regional indices

if invcon/f_ig.nonTCresp
   disp('WARNING: compilation of cyclostationary response DIFFERS from TransCom !')
end

c = 1:monperyr
% PRESUBS
greens(c,:,itrcp) = squeeze(tspresub(c,ipre,:))' ...
   + squeeze(tspresub(c+monperyr,ipre,:))' ...
   + squeeze(tspresub(c+2*monperyr,ipre,:))' ...) − (2.0 + fracyr(1,c))*ppmv*ones(nsites,1)*presubGTC(ipre);

% PULSES (land/ocean)
for i = 1:mpulse % c represents the elapsed months after the pulse in month i
   actm = i + c − 1; % actual month: this represents the month for which the response is written. Depends on the month of pulse (i) and the month’s number we are in the cyclostationary year (c).
   % Example A: i=2 (pulse was in February), c=2 (2. month after pulse)
   % Then mresp=3 (March), because Feb: 1. month, Mar: 2.month. % Example B: i=9 (pulse was in September), c=6 (6. month after pulse).
   % Then mresp=2 (February), because Sep: 1. month, ..., Feb: 6.month.
   if actm > monperyr
      actm = actm − monperyr;
   end
   % assign regional pulses to tracers
   itrcs = i + (ireg − 1).*mpulse;
   % calculate for all sites the cyclostationary response to
   % a pulse in month actm from all regions
   if invcon/f_ig.nonTCresp
      greens(1:actm,:,itrcs) = squeeze(tsmon(c,1:actm,ireg,:))' ...
         + squeeze(tsmon(c+monperyr,1:actm,ireg,:))' ...
         + squeeze(tsmon(c+2*monperyr,1:actm,ireg,:))' ...
         − 2.0*fracyr(1,c)*ppmv*ones(nsites,1)*presubGTC(1:actm);
   else
      greens(1:actm,:,itrcs) = squeeze(tsmon(c,1:actm,ireg,:))' ...
         + squeeze(tsmon(c+monperyr,1:actm,ireg,:))' ...
         + squeeze(tsmon(c+2*monperyr,1:actm,ireg,:))' ...
         − (2.0 + fracyr(1,c))*ppmv*lenfrac(c);
   end
end

clear tsmon tspresub
% **********************************************************************% build Frechet matrix % **********************************************************************
% The Frechet matrix is just the reshaped Green's array:
frech = reshape(greens,ndat,nsources);
% append constraints to Frechet matrix 
% annual ocean OIP results if invcon/f_ig.sepOIPeqs
idxoffset = monperyr * invcon/f_ig.nlandreg; % skip land sources
for r = 1:invcon/f_ig.noceanreg
   lastrow = zeros(1,nsources);
   midx = ((r−1)*monperyr + 1):(r*monperyr);
   lastrow(midx+idxoffset) = 1;
   frech = [frech; lastrow]; %#ok<AGROW>
end
% neutral Trop+South land constraint (annually) if invcon/f_ig.TSland0
% the aggregated TSland region consists of % the regions [3,4,5,6,9,10] with all months
lastrow = zeros(1,nsources);
TSidx = [3,4,5,6,9,10]; % regional indices
for i = 1:length(TSidx)
   r = TSidx(i); % regional index for Trop America
   midx = ((r−1)*monperyr + 1):(r*monperyr); % monthly indices
   lastrow(midx) = 1;
end
frech = [frech; lastrow];
% Trop. America constraint based on RAINFOR (annually; Phillips [2009])
if invcon/f_ig.RAINFOR
   r = 3; % regional index for Trop America
   midx = ((r−1)*monperyr + 1):(r*monperyr);
   % monthly indices
   lastrow(midx) = 1;
   frech = [frech; lastrow];
end
% 4 annual land constraints after Sarmiento, 2010
% a)Trop America: region 3
% b)Trop Asia plus Trop Africa: regions 5,9 % c)Eurasia (temp.+boreal): regions 7,8,11 % d)N.America (temp.+boreal): regions 1,2 if invcon/f_ig.Sarmiento
% regr = [3,5,9,11]; % regional indices
% regr = [3,5,9,11]; % regional indices
% regr = [3,5,9,11]; % regional indices
% regr = [3,5,9,11]; % regional indices
Appendix A. Cyclostationary CO$_2$ response fields

 ireg = [7,8,11]; % regional indices
 for i = 1:length(ireg)
   r = ireg(i);
   midx = ((r-1)*monperyr + 1):(r*monperyr); % monthly indices
   lastrows(3,midx) = 1;
 end

 ireg = [1,2]; % regional indices
 for i = 1:length(ireg)
   r = ireg(i);
   midx = ((r-1)*monperyr + 1):(r*monperyr); % monthly indices
   lastrows(4,midx) = 1;
 end

 frech = [frech; lastrows];

 seasonal smoothing for tropical land regions
 −Trop America, region 3
 −Trop Africa, region 5
 −Trop Asia, region 9
 −South Africa, region 6 (added on 3.4.2011 by KS)

 if invcon/ig.smoothTL > 0
   lastrows = zeros(12*3,nsources);     ireg = [3,5,6,9]; % regional indices
   for j = 1:length(ireg)
     r = ireg(j);
     midx = ((r-1)*monperyr + 1):(r*monperyr); % monthly indices
     for i = 1:11
       lastrows(i+(j-1)*monperyr,midx(i:i+1)) = [1 −1];
     end
     lastrows(12+(j-1)*monperyr,midx([12 1])) = [1 −1];
   end
   frech = [frech; lastrows];
 end

 clear lastrows* midx idxoffset TS*idx TA*idx ireg

 % save Green’s array and/or Frechet matrix in file if desired
 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
 greendumpfile = strcat(outconfig,outpath,invconfig.greendumpfile,'_',
 greenname,’.dat’);
 frechdumpfile = strcat(outconfig,outpath,invconfig.frechdumpfile,’_',
 greenname,’.dat’);

 if invconfig.dumpgreens
   f95 = fopen(greendumpfile,’wt’);
   form = {str2num(’%7.3f’,nsites),’\n’};
   fprintf(f95,’%s\n’,’Green’s functions’);
   fprintf(f95,’%s\n’,’\n’);
   fprintf(f95,’%s\n’,’(monperyr; nsites; nsources)’);
   for n = 1:nsources
     fprintf(f95,’%s\n’,’Greens(c,n)’);
   end
   fclose(f95);
 end

 if invconfig.dumpfrechet
   f96 = fopen(frechdumpfile,’wt’);

   form = {str2num(’%8.4f’,nsources),’\n’};
   fprintf(f96,’%s\n’,’Frechet/Jacobi matrix’);
   fprintf(f96,’%s\n’,’(neqns; nsources)’);
   fprintf(f96,form,frech(n,:));
   fclose(f96);
 end
Appendix B

Main Matlab routine and configuration file for the joint inversion

Here I provide the commented source code of the main routine of my joint ocean-atmosphere inversion (figures B.1 and B.2). After setting up the environment and defining some global parameters, the routine reads all necessary input from the configuration file. The version of the configuration file used for the control inversion in chapter 4 is shown in figure B.3. The routine then calls the subroutine "cycloinv" which performs the cyclostationary inversion by calling several subsubroutines. From the posterior PDF the model mean reduced $\chi^2$ is computed. The observational uncertainties are scaled iteratively until $\chi^2 \approx 1$ is achieved.
### Appendix B. Main Matlab routine and configuration file

**doinv.m**

% This routine repeatedly calls the cyclostationary inversion route, scaling each time the observational uncertainties, until the model mean chi^2 value is close to 1.

% define local environment
% add paths to access inversion code
path(path,'/home/kay/atmoinv/cycloinv');
path(path,'/home/kay/atmoinv/common'); path(path,'/home/kay/atmoinv/cycloinv/postprocessing'); path(path,'/home/kay/atmoinv/OIP_/f_lux_estimates');
% add supplemental Matlab paths
path(path,'/home/kay/Matlab'); % set absolute path to configuration file
con_file = '/home/kay/atmoinv/cycloinv/cyclo.config';
fclose('all');
% set up a few global variables
% annual total of pre−subtracted tracers in PgC/yr
presubGTC = [5.8116, 6.1729, 7.682e−4, −2.1918];
% optimize chi^2 parameter
nstep = int32(0);
while (mchi2 < 0.99 || mchi2 > 1.01) && (nstep < 10)
    nstep = nstep + 1;
    % get the control parameters
    disp('Start of inversion procedure ———');
    [invconfig, obsconfig, modconfig, outconfig, grpconfig] = ...
        read_config(con_file);
% go to inversion directory
cd(invconfig.workdir); %#ok<MCCD>
% call cyclo−inversion routine
% This is to bypass the pCOLunc entry in the configuration file:
Cologn = [5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.75 1.5 1.25 1.0 0.75 0.5 0.25 0.1];
for i = 1:length(Cologn)

### Figure B.1

Main Matlab routine for the joint inversion, part 1 of 2.
% run inversion again
[data, sources, groups] = cycloinv(invcon/f_ig, obscon/f_ig, ...
    modcon/f_ig, outcon/f_ig, grpcon/f_ig);

else
% mchi2 too large, ie. data unc. should be increased
% Note that the following adjustment rule is purely intuitive!
if mchi2 >= 4.0
    error('doinv: chi2 too large (>= 4.0) !')
elseif mchi2 >= 2.0
    obscon/f_ig.scale = obscon/f_ig.scale * (1.0 - 0.3*(mchi2 - 1.0));
else
    obscon/f_ig.scale = obscon/f_ig.scale * (1.0 - (mchi2 - 1.0));
end
% run inversion again
[data, sources, groups] = cycloinv(invcon/f_ig, obscon/f_ig, ...
    modcon/f_ig, outcon/f_ig, grpcon/f_ig);

% calculate model mean chi^2 value as before
srcupdate = sources.postest − in/f_lateND(sources.priorest,nAGCM);
mismatch = data.postdat − in/f_lateND(data.priordat,nAGCM);
sumsrc = squeeze(sum(srcupdate.^2.0./in/f_lateND(diag3D(sources.priorcov),
    nAGCM),1));
sumdat = squeeze(sum(mismatch.^2.0./in/f_lateND(data.priorvar,nAGCM),1));
sumdat = sumdat/data.ndat;
sumsrc = sumsrc/data.ndat;
mchi2 = average(sumdat + sumsrc);
disp(' ');
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
fprintf('Model mean chi^2 in current iteration: %4.2f
',mchi2);
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
disp(' ');
clear srcupdate mismatch

if nstep == 10
    disp(' ');
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
fprintf('Inversion optimization failed (%g iterations)
',nstep);  % nstep 10
fprintf('Model mean chi^2 in last iteration: %4.2f
',mchi2);
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
disp(' '); return
else
    disp(' ');
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
fprintf('Inversion optimization successful after %g iterations
',nstep);
   % nstep 10
fprintf('Model mean chi^2 in optimized inversion: %4.2f
',mchi2);
    % nstep 10
fprintf('Adjustable parameter in "optimal" inversion: %4.1f
',obscon/f_ig.
    scale);
disp('−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−');
disp(' '); end
    % nstep 10

disp(' ');
disp('−−−−−−−−−−−− Inversion procedure complete −−−−−−−−−−−−')

disp(' ') end
Appendix B. Main Matlab routine and configuration file

Figure B.3
Configuration file for the control (mode 4) inversion in chapter 4.

---

configuration file for cyclostationary inversion

allowed parameter types:
0: logical
1: integer
2: double
3: string / array
4: twin string

inversion setup:
400 cyclosteady = true.
0 pCOLinv = false.
0 joinOIP = true.
0 sepOIPreqs = true.
0 dsoind/VOGM = true.
0 nolandprior = true.
0 nooceanprior = false.
0 TSland0 = false.
0 RAINFOR = false.
0 Sarmento = false.
0 nonTcred = true.
0 testIter = false.
0 Tconstraint = false.
2 smoothTL = 10.0
0 newResponses = false.
3 title = 'cyclo-stationary'
3 workdir = '/home/kay/atmo/inv/cycloinv/'
3 sourcef = '/files4run/priors.L1and0.3.2Tak.nc'
3 Ffossilf = '/files4run/global.FossilEmissions.1751_2005.dat'
3 regmon97f = '/files4run/regional_monthly_tak97_/files_luxes.dat'
3 regmon08f = '/files4run/regional_monthly_tak08_/files_luxes.dat'
3 regareaf = '/files4run/area_TransCom3_regions.dat'
3 /filesxmatf = 'OIP_priors/Tak09_seasPatt.mat'
3 pCOLpath = 'pseudoCOLdata/
3 OIPpath = '/home/kay/atmo/inv/OIP_/files_lux_estimates/contemporary/
3 period = [2000;2008]
1 nlandreg = 11
1 noceanreg = 11
1 npresub = 4
1 onoffset = 1
1 nconst = 1
0 dumpgroups = false.
0 dumpgreen = false.
0 dumphet = false.
3 greenump = 'greensump'
3 frechump = 'frechetump'
0 deference = true.
0 shy = true.
0 verychatty = false.

handling of observational data:
193 GVpath = '/home/kay/data/GLOBALVIEW−CO2/
3 GVversion = 'gv99'
3 statlocpath = '/home/kay/atmo/inv/data/observations/
3 /filesixedStats = 'statlocs_respindex.gv99./filesixedNetwork90−08.dat'
3 T3statlist = '/files4run/list_dat75_adaptedT3L2.dat'
3 writepath = '/home/kay/atmo/inv/OIP_/files_lux_fluxes/contemporary/
3 period = [2000;2008]
1 nlandreg = 11
1 noceanreg = 11
1 npresub = 4
1 onoffset = 1
1 nconst = 1
0 dumpgroups = false.
0 dumpgreen = false.
0 dumphet = false.
3 greenump = 'greensump'
3 frechump = 'frechetump'
0 deference = true.
0 shy = true.
0 verychatty = false.

---

transport model selection (Stephens=3,4,12):
12 0 CSU−gurney = true.
0 GISS−gurney = true.
0 GISS−prothero = true.
0 JMA−CDTM−maki = true.
0 MATCH−bradshaw = true.
0 MATCH−chen = true.
0 MATCH−law = true.
0 NIES−makishima = true.
0 NIES−taguchi = true.
0 PCTM−zh = false.
0 TM2−sice = true.
0 TM3−heimann = true.

output configuration:
5 3 output = '/home/kay/atmo/inv/cycloinv/output/
0 dsores = true.
0 dgroups = true.
0 dodata = true.
0 donetcdf = false.

declaration of explicit groups:
4 4 'ind.north' = 'files4run/groups/ind.north'
4 'ind.south' = 'files4run/groups/ind.south'
4 'ocn.north' = 'files4run/groups/ocn.north'
4 'ocn.south' = 'files4run/groups/ocn.south'
4 'ind.all' = 'files4run/groups/ind.all'
4 'ocn.all' = 'files4run/groups/ocn.all'
3 types = {'annual';'aggannual';'aggmonth';'globmonth'}
Appendix C

Curriculum Vitae (CV)

Kay Steinkamp

Inst. f. Biogeochemistry and Pollutant Dynamics
CHN E 28.1
Universitaetstr. 16
CH - 8092 Zurich
Switzerland
Phone: +41 44 632 08 56
kay.steinkamp@env.ethz.ch

EDUCATION

Swiss Federal Institute of Technology (ETH) Zurich, Switzerland

University of Freiburg, Germany / Fraunhofer Institute for Solar Energy Systems (ISE)
Major in Complex systems and Astrophysics.

University of Karlsruhe, Germany

SCIENTIFIC EXPERIENCE

ETH, Zurich, Switzerland, 6/2007 - 9/2011
Graduate Student

- Development of a seasonal joint ocean-atmosphere inversion for the estimation of regional sources and sinks of atmospheric CO$_2$.

- Compilation of multiple data streams to constrain the joint inversion, including concentrations of atmospheric CO$_2$, ocean interior DIC and surface ocean pCO$_2$ as well as information on terrestrial CO$_2$ fluxes and atmospheric transport models.

- Estimation of cyclostationary CO$_2$ sinks and sources for the last three decades (1980-2008).

- Assessment of the potential of the inclusion of remotely sensed CO$_2$ column data in the atmosphere.

- Research visit to IFM-GEOMAR on the inversion of N$_2$O observations, Kiel, Germany.

Diploma Student and Research Assistant

- Development of a non-isothermal PEM fuel cell model with special focus on water transport (in gaseous and liquid phases) mechanisms in the membrane.

- Dynamic simulations with the fully coupled PEM model on timescales of seconds.

- Insights into fuel cell engineering and operation.

COMPUTER SKILLS
MS Windows/Office, Linux, Mac OS, C/C++, FEMLab, Matlab, Mathematica, IDL, Fortran, LaTeX, Ferret, Adobe CS, Photoshop.

TEACHING & MENTORING

Teaching assistant for Dr. M. Munnich, *Aquatic Physics*, ETH Zurich, 2009-2010.

Teaching assistant during the Synthesetage, EAWAG aquatic research, Dübendorf, Switzerland, 2008.


Tutor for Mathematics for engineers and computer scientists II, University of Freiburg, Germany, 2006.

Tutor for Mathematics for engineers and physicists I+II, University of Freiburg, Germany, 2003-2004.

SCHOOLS


EUR-OCEANS/CARBOOCEAN Summer School on data assimilation in biogeochemical oceanography, Kiel, Germany, 9/2007

PUBLICATIONS


PRESENTATIONS & POSTERS


Steinkamp K. and Gruber N., *Constraints on air-sea CO₂ fluxes from the ocean perspective*, Annual TransCom meeting, Jena, Germany, 2009 (presentation).


**GRANTS**

**Ph.D.**
- COST fellow, 2007-2010
- CARBONES fellow, 2010-present

**ADVISORS & RESEARCH COMMITTEE MEMBERS**

**Ph.D.**
- Prof. Dr. Nicolas Gruber Guyan (ETH), Prof. Dr. Sonia Seneviratne (ETH), Dr. Philippe Peylin (LSCE, Paris, France).

**Diploma**
- Prof. Dr. Joachim Luther (ISE), Dr. Jürgen Schumacher (CCP ZHWin, Winterthur, Switzerland), Prof. Dr. Mario Ohlberger (Uni Freiburg), Dr. Christoph Ziegler (ISE).
List of Tables

2.1 Gas transfer parameterizations from the literature. ............................................. 48

2.2 Overview on selected spaceborne instruments for CO$_2$ column retrieval. ........ 58

3.1 Annual carbon fluxes for aggregated land regions used to constrain the joint inversion. ......................................................................................................................... 71

3.2 Annual posterior fluxes for all 22 regions in the inversion as well as some additional large-scale regions. Results from all inversion modes are compared to the studies of Gurney et al. (2004) (T3L2) and Jacobson et al. (2007b) (Ja07). Bold numbers indicate 90% statistical significance for the flux sign. ... 75

4.1 Ocean and atmosphere transport models, for which Green's functions (footprints) were used in this study. In bold: Stephens’ model subset. ......................... 109

4.2 Annual posterior fluxes for all 22 regions in the inversion as well as selected aggregated regions. Control results (mode 3, common network) are compared to the study of Gurney and Eckels (2011) (GE). Bold numbers indicate 90% statistical significance for the flux sign. Highlighted cells contain fluxes that differ by more than 0.5 Pg C yr$^{-1}$ from the individual network inversion (see table in appendix for more details). ................................................................. 111

4.3 Annual posterior fluxes for all 22 regions in the inversion as well as selected aggregated regions. Control results (mode 3, common network) are compared to the study of Gurney and Eckels (2011) (GE), as well as to the results from all modes and both networks. Bold numbers indicate 90% statistical significance for the flux sign. Explicit uncertainties are omitted for clarity here, but can be found in table 4.4. ................................................................. 132

4.4 Annual posterior total uncertainties for all fluxes in the inversion as well as selected aggregated regions. Flux values are tabulated in table 4.3. ............... 133
List of Figures

1.1 Atmospheric carbon dioxide concentration over the past millennium. Until the onset of the industrial revolution around 1750 the CO$_2$ levels had been stable at about 280 parts per million (ppm), since then they increased by more than 38% (to 387 ppm in 2009). Data before 1958 are based on measurements of air bubbles trapped in ice cores from Antarctica. Later measurements (see also inset) were taken at Mauna Loa, Hawaii. They resolve the seasonal cycle imposed by the terrestrial biosphere in the Northern Hemisphere and show CO$_2$ drawdown in spring and summer due to photosynthetic productivity and growth as well as elevated CO$_2$ levels in fall and winter when respiratory processes dominate. Figure based on Sarmiento and Gruber (2002). 2

1.2 Sketch of the Earth’s annual and global mean energy balance. On longterm average, the amount of incoming solar radiation absorbed by the Earth and atmosphere is balanced by the release of the same amount of outgoing longwave radiation. About half of the incoming solar energy is absorbed by the Earth’s surface and subsequently transferred back to the atmosphere by warming the air in contact with the surface (thermals), by evapotranspiration and by longwave radiation that is absorbed by clouds and greenhouse gases. The atmosphere then re-emits this longwave energy into all directions, including out to space. Figure from Kiehl and Trenberth (1997). 5

1.3 Global carbon cycle showing carbon stocks and fluxes between the three main compartments: the atmosphere, the ocean and the land. Anthropogenic contributions are in red; natural fluxes in black. Numbers represent an approximation for the 1990s average. NPP means net primary production. Figure from Sarmiento and Gruber (2002), numbers updated according to IPCC (2007a). 7
1.4 Major fluxes in the global carbon cycle over (roughly) the Anthropocene. Each flux contribution corresponds to a term in the global CO$_2$ budget (1.1). Anthropogenic CO$_2$ emissions are shown as positive fluxes into the atmosphere; they comprise emissions from fossil fuel combustion and other industrial processes (such as cement manufacture and gas flaring; $F_{Foss}$) and emissions from land-use change (mainly deforestation and after the mid-1980s mainly tropical deforestation; $F_{LUC}$). Negative fluxes denote the accumulation of carbon in the three major compartments: the atmosphere ($dC_a/dt$), the land biosphere ($F_{LandSink}$) and the ocean ($F_{OceanSink}$). Numbers in the right panel give the 2000-2008 average estimates from Le Quere et al. (2009). Figure of Rau-pach and Canadell (2010), based on Le Quere et al. (2009). 13

1.5 Global distribution of fossil fuel emissions for the year 1995. The color scale is logarithmic. The right panel shows the zonally integrated fossil fuel emissions. The emerging pattern demonstrates the dominant fossil fuel sources from industrial countries in temperated northern hemisphere regions. The figure is based on data from Brenkert et al. (1998). 15

2.1 Minimization of a quadratic cost function $J(x)$, defined over a two-dimensional parameter space. The transition from the prior estimate $x_0$ to the optimized posterior estimate $x_{opt}$ is indicated. The projections of the cost function on the parameter plane are elliptic iso-cost lines with semi-axis lengths determined by the gradient of the cost function and related to the uncertainty and error covariance of the optimal parameter estimate. 31

2.2 Shown are in blue and green colors the 11 regions over ocean and land, respectively. Red points show the observation sites for atmospheric CO$_2$. White points correspond to the locations of ocean interior DIC profile measurements used in the ocean inversion underlying mode 2 (and all further modes) of our joint inversion. For the 1992-1996 inversion, monthly mean atmospheric CO$_2$ measurements from 85 sites and more than 60,000 ocean DIC observations are used. The decadal mean inversions use various atmospheric CO$_2$ networks, see figure 4.2 in chapter 4 for details. 33
2.3 Global distribution of NEP (Net Ecosystem Production) for land regions and net carbon flux across the air-sea interface (positive fluxes are directed into the atmosphere). The upper panel shows the distribution averaged over the boreal winter (here October to March), while boreal summer (here April to September) is shown in the lower panel. Air-sea fluxes are on the same scale as terrestrial biospheric fluxes to point out the terrestrial dominance in the seasonal signal. These distributions are used in the synthesis inversion described in the text to infer regional flux patterns. The seasonal biospheric exchange is based on the Carnegie Ames Stanford Approach (CASA) model (Randerson et al., 1997) and has an annual flux of zero. The oceanic exchange is based on pCO$_2$ measurements of Takahashi et al. (1997).

2.4 Global zonal mean response functions or footprints for the 11 Transcom oceanic regions. The colored pattern is the quasi-steady state dye tracer concentration in response to the injection of a continuous flux of 1 Pg C yr$^{-1}$ into each region. Units are arbitrary, only the spatial tracer distribution matters for the inversion. Black dots represent measurement locations (only schematically), where the response functions are sampled; these “discretized” responses form the ocean transport matrix $M_o$ in the cost function.

2.5 Concentration profiles of ocean carbon tracers along North-South transects in the Pacific (left panel) and Atlantic (right panel) ocean basins. The rows represent the carbon concentration terms as defined in the text and in equation (2.23). From top to bottom they are observed DIC, the biological correction term ($\Delta C_{bio}$), the anthropogenic carbon concentration ($\Delta C_{ant}$), and the preindustrial gas exchange tracer ($\Delta C_{gasea}$). Selected isopycnal surfaces (in kg m$^{-3}$) are contoured in the bottom row. Black dots in the top row show the locations of the DIC observations from which these fields were interpolated; their latitudinal positions are very similar to the profile locations depicted in figure 2.2. Thick vertical lines represent the regional boundaries defined within the Transcom project, cf. blue regions in figure 2.2. Thin vertical lines show the boundaries from 30 regions as defined in the ocean inversion, before they were aggregated to the 22 Transcom regions. Figure from Jacobson et al. (2007b).

2.6 Dataset of Takahashi et al. (2009b). The top picture show the annual mean, net air-sea flux climatology for a specific gas exchange model, based on more than 3 million pCO$_2$ measurements globally. Measurement locations (cruise tracks) are shown below. In the lower right the monthly observational density is shown, separated into observations from the Northern and Southern Hemispheres.
2.7 Posterior fluxes, uncertainties and error correlations as estimated in the ocean inversion for selected regions. An ellipse represents the 68% confidence area for the flux estimate for each pair of regions. That is, those iso-cost lines that enclose the area where the posterior flux pair is located in with 68% probability; compare also to the projections in the conceptual drawing in figure 2.1. The cross within the ellipse gives the standard deviation for each of the fluxes. The fact that the crosses are always located in the ellipse’s interior without touching its edges is a general property of multi-variate Gaussian PDFs. The ellipse’s tilt indicates error correlation; for example, two perfectly anti-correlated flux errors would cause the ellipse to be oriented along the dashed line with slope -1. Numerical values above the ellipses give the error correlation coefficients. The top panel compares the three major ocean basins, whereas the lower panel focuses on latitudinal correlations.

2.8 Illustration of the linear Bayesian cost function (2.28). Arrows indicate where (i.e. in which cost function terms) the data streams from the atmosphere and ocean enter the inversion system. The data streams are atmospheric CO$_2$, ocean interior DIC and surface ocean pCO$_2$.

2.9 Posterior results for modes 1, 2 and 3 of the 1990-1999 decadal inversion of chapter 4 for the tropical and northern extra-tropical land regions. Gray colors represent the atmosphere-only inversion (mode 1), blue colors the joint ocean-atmosphere inversion (mode 2), and the green ellipse is the mode 3 model mean estimate. The filled ellipses represent estimates for an individual transport model (mode 1) or an individual transport model pair (ocean and atmosphere, mode 2) and their uncertainty and error correlation. The more an ellipse is tilted towards the dashed line of -1 slope, the stronger the anti-correlation is. The framed unfilled ellipses represent the model mean result, where the lengths of the semi-axes are determined by the total posterior covariance matrix. The crosses inside the model mean ellipses give one standard deviation for the marginal total probability density for both regions.

2.10 Distribution of column mean CO$_2$ data. a) preliminary retrievals from the GOSAT satellite for selected months in 2009 and 2010 in ppm, picture source: http://www.gosat.nies.go.jp/index_e.html. b) simulated column mean CO$_2$ data from the TM3 atmospheric transport model for each month in a cyclostationary year. Shown are the concentration anomalies in ppm at each model pixel, according to equation (2.19) after presubtraction of the background responses.
3.1 Black lines show the posterior results from the control (mode 4) inversion for the 11 ocean regions. According to the sign convention, a positive flux is directed into the atmosphere. Gray shaded areas represent the total posterior uncertainty as one standard deviation, which in turn includes contributions from the Bayesian "within" error as well as from the "between" error arising from transport model spread. Results are compared to the T3L2 (Transcom 3, Level 2) atmosphere-only inversion of Gurney et al. (2004) in green, for which the monthly uncertainty is not shown for clarity. In light blue (uncertainty given by the dashed line) the ocean constraint is shown, as included in joint inversion mode 2. ................................. 77

3.2 Control fluxes for the land regions are compared to results obtained in each mode of the joint inversion. ................................................................. 78

3.3 The total uncertainty (squared sum of "within" and "between" error contributions; one standard deviation; logarithmic axis) is plotted against the annual mean flux from aggregated, latitudinal land regions. Yellow colors separate global land into Northern Hemisphere extratropical (NH) land and aggregated Tropical and Southern Hemisphere Land (TSL). Green colors further partition TSL into Tropical land and Southern Hemisphere extratropical (SH) land. Gray shaded areas represent flux sign significance regimes between 70% and 99%, i.e. a point located in such a regime represents a flux that is positive (or negative) with the indicated probability. The numbers inside the symbols give the inversion mode: number 1 represents the atmosphere-only inversion, while number 4 represents the control inversion. Black lines show the trajectory of inverse results when more and more constraints are considered. .................. 80

3.4 Control fluxes for selected aggregated regions are compared to results obtained in each mode of the joint inversion. ................................. 82

3.5 Comparison of seasonal flux amplitudes from our control inversion with T3L2. ................................. 84

3.6 Black lines show the posterior results from the control (mode 4) inversion for the 11 land regions. According to the sign convention, a positive flux is directed into the atmosphere. Gray shaded areas represent the total posterior uncertainty as one standard deviation, which in turn includes contributions from the Bayesian "within" error as well as from the "between" error arising from transport model spread. Results are compared to the T3L2 (Transcom 3, Level 2) atmosphere-only inversion of Gurney et al. (2004) in green, including their reported total uncertainty. ................................. 87
3.7 Annual mean control results for selected aggregated regions. Inventory-based estimates, which are also included in mode 3 of the joint inversion, are shown in pink. Results are compared to estimates of T3L2 (Gurney et al., 2004) and Jacobson et al. (2007a,b). 89

3.8 Annual posterior flux uncertainty for a given column data uncertainty for the pseudo-inversions. Red curves represent the pseudo-inversion constrained by PBL data at existing station locations. Blue curves represent the inversion of the column pseudo data sampled at each pixel of the underlying transport model. Black curves represent the inversion of the column pseudo data sampled at only the station locations. Dashed lines show the results if only the pseudodata are used to constrain the inversion, while solid lines show the results if the ocean prior is used in addition. Yellow markers indicate important features described in the text. 90

4.1 Annual CO$_2$ uptake by global oceans and land. The uptake is calculated by difference from fossil fuel emissions and observed global growth rate in atmospheric CO$_2$. The latter was derived from time series measurements at Mauna Loa and South Pole stations. The curve represents the station mean and the shade their spread. The mean global uptake for the three inversion periods is given as dotted lines. Note that these values are based on a fit over each period as a whole (cf. methods section), and may not equal the mean of the annual values exactly. The ocean uptake represents the weighted mean result from the ocean interior inversion (see Mikaloff Fletcher et al. (2006, 2007) for details on model weighting). It is corrected for outgassing due to carbon input by rivers. The net land-atmosphere flux is the difference between the global uptake and the ocean sink. 98

4.2 Locations of atmospheric CO$_2$ stations for all networks considered in this study. Thick gray lines in the background delineate the 22 Transcom regions, consisting of 11 oceanic and 11 land regions. 100
4.3 The normalized mismatch between posterior and prior data is plotted against prior data uncertainty (1 standard deviation). The left panel shows the unconstrained inversion (mode 1), while in the right panel the ocean constraint is included (mode 2). For each data point \(i\), where \(i\) represents a combination of station and month, the absolute mismatch \((Mx_i - d_i)\) was divided by the prior data uncertainty \(\sigma_{d,i}\) to yield the normalized mismatch \((\frac{(Mx_i - d_i)}{\sigma_{d,i}})\). \(M\) indicates the transport matrix, and the vectors \(x\) and \(d\) contain the flux variables and data points, respectively. The closer a value is to zero, the smaller the mismatch is relative to the prior uncertainty and the better the inversion was able to reproduce the station data accurately. Values between -1 and 1 (within the area framed by the dotted lines) indicate a mismatch that is within one standard deviation from the prior data.

4.4 Atmospheric CO\(_2\) constraint for the three inversion periods. Shown are prior data estimates (observations) as dashed lines and posterior data estimates as solid lines. Uncertainty bars represent the observational uncertainty assigned during data processing, as explained in the method section.

4.5 Perturbation of atmospheric CO\(_2\) compared to a preindustrial value of 280 ppm is shown as dashed line. The solid curve represents the function \(\phi(t)\) used to scale the anthropogenic ocean flux, with year 2000 as a reference. See text for details.

4.6 Change in annual fluxes in response to a prescribed Southern Ocean flux of zero. Changes are based on the difference between results from the inversions mode 2 (ocean constraint) as well as mode 3 (ocean and model constraints) and the respective inversions where the Southern Ocean flux is prescribed as zero. The increasing positive flux change over the periods is caused by the fact that the Southern Ocean sink is estimated to increase over time in the normal inversions.

4.7 Partitioning of the global carbon sink in the control inversion. For each period three levels are considered: 1) partitioning of global sink into land and ocean uptake, 2) split of the global land uptake into fluxes from Northern (extratropical) and Tropical plus Southern Land (TSL) regions, and 3) individual contributions from Tropical and Southern land regions to the TSL flux. All fluxes are normalized by the period-mean global carbon sink, which is added as number below the global sink bar. Errorbars represent normalized 1\(\sigma\) total uncertainty, i.e. the quadrature of within-uncertainty and model spread.
4.8 Decadal flux changes in Pg C yr$^{-1}$ for land regions (world map) as well as latitudinal aggregates (right panel). The 1980-2008 average was subtracted to show relative changes. Trends for the 1980-2008 period were calculated by a regression and are given as numbers in each inset. In addition to the control case (green), results are shown from the unconstrained inversion on the common network (blue), and the inversion with ocean and model constraints on individual networks (dashed green). Estimates of Gurney and Eckels (2011) are added in red. Connecting lines are only shown for clarity, they do not represent interannual variation.     118

4.9 Decadal flux changes in Pg C yr$^{-1}$ for all 11 land and 11 ocean regions as well as latitudinal aggregates. The 1980-2008 average was subtracted to show relative changes. Trends for the 1980-2008 period were calculated by a regression and are given as numbers in each inset. In addition to the control case (blue), results are shown from the unconstrained inversion on the common (gray) and individual (dashed gray) network, and the inversion with ocean and model constraints on individual networks (dashed blue). Estimates of Gurney and Eckels (2011) are added in green. They did not report ocean fluxes, therefore the trends are set to NaN for all oceanic regions as well as aggregates involving oceanic regions. Results from the unconstrained inversion on individual networks are excluded from analysis due to instability issues in the 1980-89 period (see text for further information).     120

4.10 Seasonal and annual mean net CO$_2$ exchange between the atmosphere and three selected regions: a) Boreal America, b) Tropical America, and c) Tropical and Southern Land. Light blue patches represent total monthly flux uncertainty (quadrature of within-uncertainty and model spread; $1\sigma$) for the control inversion (mode 3 on the common network). The bars show the annual mean and embrace $1\sigma$ total uncertainty.     124

5.1 Schematic view of the CARBONES data assimilation system. The system comprises a land ecosystem model (ORCHIDEE), an ocean biogeochemical model (PISCES) and an atmospheric transport model (LMDZ). Parameters of all the models are optimized with regard to several observational constraints: atmospheric CO$_2$ from satellites and surface stations, direct CO$_2$ flux measurements from towers, and carbon inventories from forests and soils.     143

A.1 Matlab routine for the compilation of the model responses (Frechet matrices), part 1 of 3.     148
A.2 Matlab routine for the compilation of the model responses (Frechet matrices), part 2 of 3. .......................... 149
A.3 Matlab routine for the compilation of the model responses (Frechet matrices), part 3 of 3. .......................... 150
B.1 Main Matlab routine for the joint inversion, part 1 of 2. ......................... 152
B.2 Main Matlab routine for the joint inversion, part 2 of 2. ......................... 153
B.3 Configuration file for the control (mode 4) inversion in chapter 4. ............ 154
Bibliography


physical mechanisms that produce realistic distributions of tracers and productivity. *Global Biogeochemical Cycles*, 18(4).


Maksyutov S. and Inoue G. Vertical profiles of radon and co$_2$ simulated by the global atmospheric transport model, (2000).


