Inverse modelling of carbon dioxide surface fluxes – estimating uncertainties due to model design and observational constraints

Zur Erlangung des akademischen Grades eines
DOKTORS DER NATURWISSENSCHAFTEN
von der Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Dipl.-Phys. Arne Babenhauserheide
aus Leimen

Tag der mündlichen Prüfung:	15. Januar 2016
Referent:	Prof. Dr. Johannes Orphal
Korreferent:	Prof. Dr. Peter Braesicke

Abstract

Sources and sinks of atmospheric carbon dioxide largely control future climate. They moderate the fraction of emitted carbon which remains in the atmosphere, the main anthropogenic driver of global warming. However the sources and sinks are hard to measure directly. Therefore they are estimated using inverse models which combine a prior estimate, inferred from characteristics of biosphere and oceans, with atmospheric measurements and adjust the sources and sinks to fit the atmospheric measurements.

Deriving a robust estimate of the global distribution of sources and sinks requires estimating systematic errors. Previous studies investigated, among other parameters, the uncertainty due to atmospheric transport which connects surface fluxes to atmospheric measurements. This study is the first to investigate the uncertainty of surface fluxes due to the choice of the inverse method and the representation of fluxes using real measurements and two well-established inverse models. The inverse models are run with harmonized input, prior fluxes, measurements, atmospheric transport and flux covariance. Comparing the mismatch between the calculated atmospheric concentration and measurements which are not used to adjust the fluxes, gives an estimate of the quality of the inversion. The difference of this mismatch between the models is smaller than the uncertainty of the mismatch. Therefore differences in the fluxes estimated by the different models provide an estimate for the contribution of inter-model errors to the uncertainty of estimated sources and sinks.

For the sink in North America, where the density of measurement sites is highest, this study finds a lower limit for the uncertainty of 0.1 Pg carbon per year, about 10% of the estimated biospheric sink. For other continents the uncertainty is on the order of 0.25 Pg carbon per year. Varying the number of observation sites used in the models showed that this uncertainty is controlled by the density of measurements. Integrating additional measurements reduces the uncertainty due to differences between the models.

To investigate the effect of complementary observations, measurements of the aggregated vertical column of CO_2 from ground-based spectrometers were implemented in one of the models. Evaluating the calculated fluxes when using these additional measurements showed that total column measurements correct mismatches introduced by using temporally sparse aircraft measurements. The strength of the Eurasian biospheric sink was derived as 3.5 ± 1 Pg carbon per year, and it was shown that a robust estimate of the European sink requires at least one additional measurement site in boreal Asia.

This study completes the assessment of different contributions to the uncertainty of inverse source/sink estimates of CO_2. It shows that adding measurements decreases the uncertainty of the estimated fluxes due to differences between the models and that total column measurements complement in-situ measurements. To this end it implements usage of ground-based total column measurements for inverse modelling which lays the foundation for adding more measurement sources, from ground as well as from satellite.

Contents

1 Introduction **1**

2 Understanding, measuring and modelling CO_2 in the atmosphere **3**
2.1 CO_2 and Climate . . . 3
2.1.1 The Carbon Cycle . . . 4
2.1.2 Understanding Sources and Sinks of Carbon Dioxide . . . 5
2.1.3 Prediction and Mitigation of Climate Change . . . 8
2.2 Measuring greenhouse gases . . . 9
2.2.1 In-situ measurements . . . 9
2.2.2 Remote sensing from ground . . . 10
2.2.3 Remote Sensing from satellite . . . 12
2.2.4 Summary . . . 12
2.3 Bottom-up modelling of CO_2 surface fluxes . . . 13
2.3.1 Biosphere . . . 13
2.3.2 Oceans . . . 15
2.3.3 Fire and Fossil Fuel Fluxes . . . 16
2.4 Atmospheric Transport of CO_2 . . . 16
2.4.1 Meteorological data . . . 17
2.4.2 Transporting tracers . . . 18
2.5 Assimilation Methods . . . 18
2.5.1 Different Approaches: ensemble or variational . . . 18
2.5.2 Assimilating measurements . . . 19
2.5.3 Ensemble Method: CarbonTracker . . . 21
2.5.4 Variational Method: TM5-4DVar . . . 25
2.6 The promise of assimilating total column measurements . . . 27

3 Uncertainty of CO_2 surface fluxes constrained by in-situ observations **29**
3.1 Comparing CarbonTracker and TM5-4DVar . . . 29
3.1.1 Harmonizing the inverse methods . . . 30
3.1.2 Transport model and observation operator . . . 30
3.1.3 Background flux and initial guess . . . 31
3.1.4 Observations and observation errors . . . 31
3.2 Uncertainties due to model design . . . 33
3.2.1 A posteriori concentration fields . . . 33
3.2.2 Comparison of a posteriori surface fluxes . . . 38
3.3 Conclusions . . . 47

4 The added value of total column observations for CO_2 inverse modelling 49
4.1 Observations and Methods . . . 49
4.2 Viability of assimilating total column data . . . 50
4.3 Estimating the representativeness error . . . 53
4.4 Joint assimilation of total column and in-situ measurements in Carbontracker . . . 57
4.4.1 Cross-comparison of modelled fields and sets of measurements . . . 57
4.4.2 The effect of assimilating aircraft measurements when using a biased prior 61
4.4.3 Unequal observational pressure guides flux attribution . . . 64
4.5 A new limit for the uncertainties of Eurasian fluxes . . . 67
4.6 Summary . . . 71

5 Conclusions and Outlook 73

6 Acknowledgements 77

7 Bibliography 79

A Appendix 93
A.1 Prior flux uncertainty harmonization . . . 93
A.2 In-situ measurement sites . . . 93
A.3 Learjet flight information . . . 109
A.4 The EnSRF Toy Model . . . 111

Chapter 1

Introduction

Carbon dioxide (CO_2) is the most important greenhouse gas which is directly influenced by human activity. Anthropogenic emissions release roughly 10 Gt carbon into the atmosphere per year (Peters et al., 2013), part of which gets taken up by the biosphere and the oceans. The net sink of atmospheric CO_2 in biosphere and oceans determines the fraction of emitted CO_2 which remains in the atmosphere, the largest driver of climate change (Stocker et al., 2013, chapter 8.5.1). But sources and sinks of CO_2 are hard to measure directly.

Consequently there are large uncertainties about the quantity and strength of sources and sinks of CO_2 – in Europe uncertainty is on the same order of magnitude as the estimated biospheric sink (Reuter et al., 2014). Methods for observing the fluxes directly require either eddy covariance measurements at multiple height levels (Foken et al., 2012; Baldocchi, 2003), measurements of concentration changes in a sealed volume of air, or biometric measurements (Davidson et al., 2002; Curtis et al., 2002). But such bottom-up approaches are only representative for a given collection of vegetation types in a limited geographic area, making it hard to scale their results up to continental scale.

To address this problem, inverse flux modelling estimates regional sources and sinks by assimilating atmospheric concentration measurements. A transport model connects surface fluxes and atmospheric concentration measurements by transporting a carbon tracer on meteorological fields (like those from ECMWF, the European Centre for Medium-Range Weather Forecasts). An inverse method approximates an optimal flux distribution by minimizing the mismatch to observations as well as the deviation from a prior flux model (Rodgers, 2000). Prior flux models scale up direct flux measurements but suffer from large uncertainties.

There are two main classes of assimilation techniques for complex inversions: variational methods and ensemble methods (Lahoz et al., 2007; Lahoz and Schneider, 2014). Variational methods use algorithms inspired by the Newton method, adjusting their estimated fluxes step by step to adapt calculated atmospheric concentrations to the atmospheric measurements (e.g. Juhász and Bölöni, 2007; Gilbert and Lemaréchal, 1989; Newton, 1669). Ensemble methods approximate fluxes using many different realizations (an ensemble). The realizations are chosen to be distributed following the mean value and uncertainty of the prior fluxes. Ensemble methods then adjust each realization using atmospheric measurements. The new mean value and uncertainty for the fluxes follow from the distribution of the adjusted realizations (e.g. Whitaker and Hamill, 2002).

This study aims at better quantifying the uncertainties involved in inverse modelling of sources and sinks of carbon dioxide. In particular, this thesis assesses the uncertainty due to the choice of the data assimilation system and due to the choice of the type of observational constraint.

It shows that these uncertainties can be reduced by combining in-situ observations and remote sensing measurements.

Comparing different inverse methods allows estimating uncertainties which cannot be captured by their intrinsic error propagation methods themselves. This study uses the methods implemented in TM5-4DVar (Basu et al., 2013) and CarbonTracker (Peters et al., 2007; Peters et al., 2005; Whitaker and Hamill, 2002) and harmonizes their input values to estimate the differences due to their setup of the state to be optimized and the inverse method used. Both methods build on the same transport model (TM5, Krol et al., 2005). Uncertainties due to the choice of transport model have previously been assessed by Gurney et al. (2004).

Currently both compared inverse methods use in-situ measurements (obspack, Masarie et al., 2014). In-situ measurements provide accurate information about fluxes at a given point with errors around 0.1 ppm (Andrews et al., 2013) but they are not very representative for model scales. The model operates on a grid with boxes of 6×4 degrees or 1×1 degrees - about 500km or 100km side-length, therefore there are many local biosphere phenomena which affect in-situ measurements but cannot be captured on the model scale. This error is commonly referred to as representativeness error.

In the latter part of the thesis, assimilation of total column measurements is added to CarbonTracker. Total column measurements contain information about more distant regions whose CO_2 signal is already diluted vertically. Since the spatial variability of the aggregated CO_2 concentration in a vertical column is about a magnitude lower than near the ground, total column measurements are more representative for the model grid scale than flask measurements from the ground, so they have a smaller representativeness error.

Ground sites for total column measurements provide high frequency data of an aggregated column of CO_2 from the ground to the top of the atmosphere. The Total Carbon Column Observing Network (TCCON, Toon et al., 2009a) provides measurement an 1σ uncertainty of 0.2% (Toon et al., 2009b) corresponding to an absolute uncertainty between 1 ppm and 0.4 ppm CO_2. There are only few total column stations, though, 18 worldwide in 2015. Different from the in-situ network which has the highest density in the USA, the density of TCCON sites is highest in Europe, so they complement the in-situ network in terms of spatial coverage.

Satellite measurements provide low frequency data with higher uncertainty, but with good spatial coverage (Butz et al., 2011; Butz et al., 2013; Basu et al., 2013). With these characteristics they are orthogonal to ground measurements and provide information which the other methods cannot deliver, especially for inaccessible regions. Satellites with CO_2 measurements suitable for assimilation include GOSAT and the upcoming OCO-2 (Dils et al., 2014; Hammerling et al., 2012).

Together, flask measurements, ground-based total column measurements and satellite total column measurements provide better constraints for the inverse methods than flask measurements alone, which reduces the uncertainty of estimated fluxes.

In my thesis I estimate the uncertainties of the calculated distribution and strength of sources and sinks of atmospheric carbon dioxide with two approaches: comparing results calculated by different inversion methods (chapter 3) and comparing results from different types of measurements (chapter 4). I evaluate the flux uncertainties from the perspective of carbon cycle investigation such as the continental-scale distribution of the biospheric sink.

Chapter 2

Understanding, measuring and modelling CO_2 in the atmosphere

2.1 CO_2 and Climate

Greenhouse gases have an important effect on climate, because they impact the energy balance between the outgoing radiation from the Earth's surface and incoming sunlight. They are the base of the greenhouse effect. Human activity predominantly affects the greenhouse effect by increasing the atmospheric concentration of CO_2 and other well-mixed greenhouse gases (GHG) such as methane, nitrous oxide (Schmidt et al., 2010).

Without any atmosphere, the temperature of the earth would be given by the equilibrium between directed incoming light in the short wave spectral range from the sun and undirected outgoing light in the infra-red which is emitted from the earth (IPCC, 2007). Energy radiated from a black body scales with the fourth power of the temperature (as first shown by Štefan, 1879; Boltzmann, 1884). Without absorption of outgoing radiation within the atmosphere due the greenhouse effect, the global mean near-surface temperature would be about $-18°$ Celsius. Absorption in the atmosphere traps radiation, raising the equilibrium temperature to the measured 15° Celsius (IPCC Working Group I, 1996a; Jones et al., 1999; Rayner et al., 2006).

This effect was already identified by Arrhenius (1896). In particular Arrhenius (1896) predicted that a doubling of the concentration of CO_2 would lead to an increase in the mean temperature by 4.9°C to 6.1°C — which is close to the 2.1°C to 4.7°C range of results calculated by modern climate models, as summarized by Flato et al. (2013, table 9.5).

Other greenhouse gases include water vapour (the strongest GHG) and ozone (Ramanathan and Coakley, 1978; Kiehl and Trenberth, 1997; Schmidt et al., 2010) as well as methane (Myhre et al., 2013, p. 698), and, to a lesser degree, the very abundant oxygen and nitrogen (Höpfner et al., 2012).

But different from many other factors, the CO_2 concentration in the atmosphere is influenced directly by human activity. The strongest contribution to the increasing CO_2 concentration are burning of fossil fuel, deforestation and cement manufacturing which introduce previously bound carbon into the atmosphere (Worrell et al., 2001; Le Quéré et al., 2015).

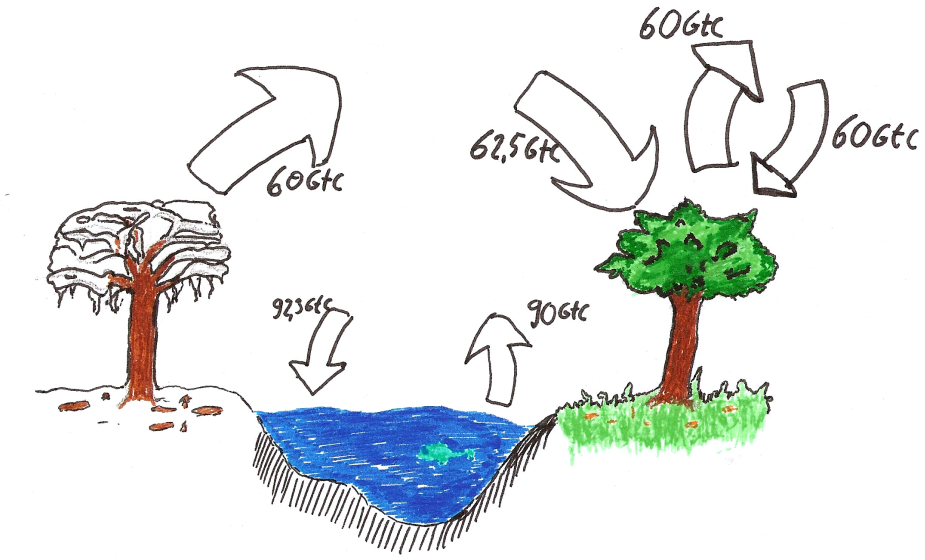


Figure 2.1: Main pathways for carbon-exchange between biosphere, oceans and atmosphere. In autumn and winter the biosphere releases most of the carbon it captured in summer, while the upper ocean keeps an equilibrium with the CO_2 concentration in the atmospheric boundary layer. Data from Le Quéré et al. (2012).

2.1.1 The Carbon Cycle

The carbon cycle is the flow of carbon between different reservoirs of carbon. It is split in a fast cycle and a slow cycle. The fast cycle exchanges carbon between the atmosphere, the biosphere and the surface ocean. It is relevant for time scales on the order of weeks to years. The slow cycle is concerned with carbon exchanges including the deep ocean and the Earth's crust. It is relevant for time scales on the order of centuries to millennia. This study is only concerned with the fast carbon cycle.

Carbon in the atmosphere mainly exists in the form of CO_2. It moves to the biosphere by photosynthesis and gets integrated by plants as sugar-compounds. Transport from the biosphere to the atmosphere mostly happens through respiration, decay and fires, which both release the carbon from the plants. Due to this effect, the mass of carbon in the atmosphere and in the biosphere is in an equilibrium. Increased partial pressure of CO_2 in the air increases the growth rate of many plants (Griffin et al., 2001), so if the CO_2 in the atmosphere rises, part of it gets taken up by the biosphere. The amount of carbon exchanged between atmosphere and biosphere per year is on the order of 120 gigatons carbon (U.S. Department of Energy Office of Science, 2008). Transport between ocean and atmosphere on the other hand mostly happens through dissolution of CO_2 in the ocean, which then gets stored in form of several different carbon-compounds (Rhein et al., 2013). The dissolved CO_2 in the ocean (DIC: dissolved inorganic carbon) and the CO_2 in the air form an equilibrium. Increased partial pressure of CO_2 in the air leads to higher uptake of CO_2 in the ocean and vice versa. The yearly exchange between atmosphere and ocean is on the order of 90 Gt per year (Le Quéré et al., 2012). This exchange is added up over the year, with large take-up of carbon by the biosphere during the summer growing season and net release

of carbon from the biosphere in autumn and winter (see Figure 2.1). Without additional sources the long term net flux between those three reservoirs would be roughly zero.

The fourth reservoir, the Earth's crust, is only relevant for the equilibrium between the reservoirs on time scales on the order of thousands of years, because its interactions with the other reservoirs are very slow. The sole exception is burning of fossil fuel, which releases about 8.3 Gt carbon per year (Le Quéré et al., 2012) - and different from interactions between the other reservoirs, this release is one-way: The carbon released from burning fossil fuels does not return to the crust in the span of a human life.[1]

Not all of that anthropogenic carbon stays in the atmosphere, though. As shown in Figure 2.2 about half of it ends up in the biosphere and ocean (Le Quéré et al., 2012; Ciais et al., 2013), so that only the other half of the carbon released through burning of fossil fuel actually remains in the atmosphere and contributes to the greenhouse effect.

This fraction which remains in the atmosphere can be measured from the average yearly increase of carbon dioxide of 2 ppm (corresponding to 4.28 Gt carbon).[2] Contrasting this increase with the 8.3 Gt carbon released from human activity provides the mean take-up from the biosphere and the oceans. The spatial distribution of sources and sinks can be estimated roughly with ocean and biosphere models which calculate the take-up of CO_2 at the Earth's surface by scaling up ecological variables like the type of plants in a given region and meteorology like the amount of sunlight which reaches the surface.

But this up-scaling is impeded by heterogeneities of the surface. So these models can only provide uncertainties on the scale of the actual surface fluxes, and their calculation can only use knowledge about the current behaviour of the biosphere, which may not fit the behaviour under climate change. An example for the impact of this uncertainty is the research by Wolkovich et al. (2012) who showed that "warming experiments underpredict advances in the timing of flowering and leafing by 8.5-fold and 4.0-fold"

To predict the future increase of the greenhouse effect from anthropogenic emissions, given these sources of uncertainty, and to detect changes in the sources and sinks of carbon, it is therefore necessary to use additional information to estimate the transport between atmosphere, biosphere and ocean – to understand the sources and sinks of carbon dioxide.

2.1.2 Understanding Sources and Sinks of Carbon Dioxide

Due to the large amount of carbon exchanged between the carbon reservoirs over the year, it is hard to estimate the net uptake and release from the biosphere using models of the biosphere and direct measurements of carbon fluxes. The total amount of carbon exchanged between atmosphere, biosphere and ocean per year is on the order of 210 Gt carbon (IPCC, 2007), while the net uptake is only on the order of 5 Gt per year, 2.5% of the total flux (Le Quéré et al., 2012; Schulze, 2006).

[1] The amount of carbon in the deep ocean is about 38 thousand gigatons. It is increased by approximately 1 gigaton per year, or 0.0026%. The sedimentation is on the order of 0.2 gigatons carbon per year. Assuming a linear relationship between carbon content in the ocean and transport into the crust, the transport should only increase by 0.0026% per year, which amounts to about 5 tons. If the deep ocean were to take up all the carbon humans released in the last 100 years, transporting 1 gigaton back into the crust would still take time on the order of half a millennium (IPCC, 2007). Therefore the equilibrium with the deep ocean is more relevant on shorter time scales than the transport to the crust.

[2] The conversion factor from ppmv CO_2 to GtC is 2.14, calculated from the molar mass of roughly $M_{CO_2} = 44g/mol$ for carbon dioxide, the molar mass of $M_C = 12g/mol$ for carbon, $M_{air} = 28.9g/mol$ for air (Halliday et al., 2003) and the total mass of the air $m_{air} 5.15 \times 10^6 Gt$ (Trenberth and Smith, 2005):
$\left(\frac{M_{air}}{M_{CO_2}} \cdot \frac{M_{CO_2}}{M_C} \cdot \frac{1}{m_{air}}\right)^{-1}$

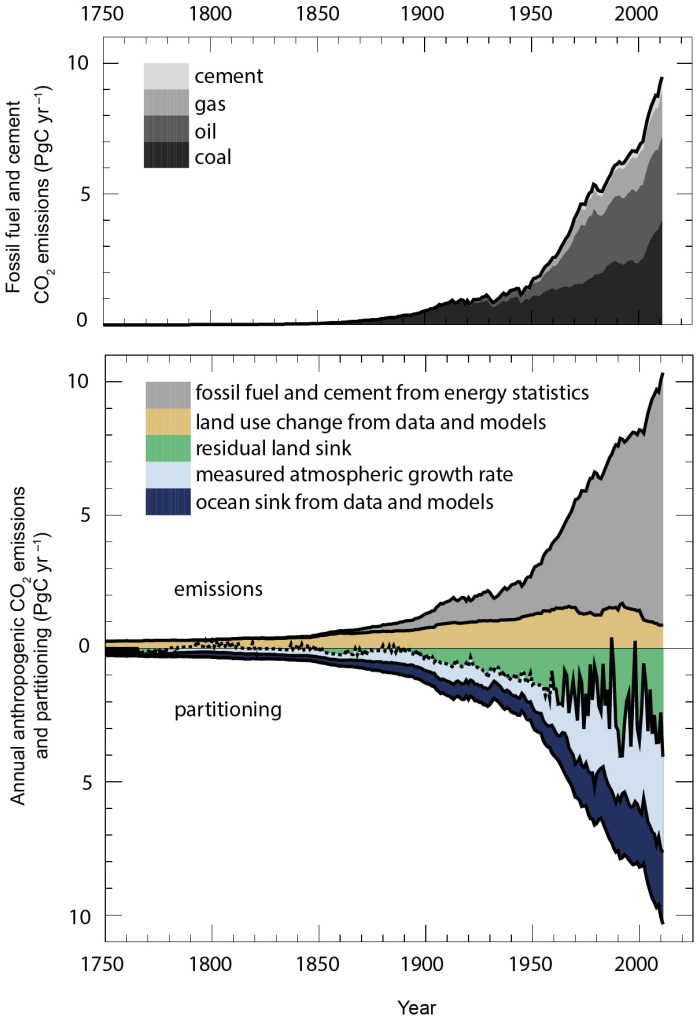


Figure 2.2: Historical annual anthropogenic CO_2 emissions and their partitioning among the atmosphere, land and ocean (PgC yr–1) from 1750 to 2011. (Top) Fossil fuel and cement $CO_2)emissionsbycategory.(Bottom)Fossilfuelandcement CO_2$ emissions as above with the addition of the repository which received the emitted carbon. The residual land sink (term in green in the figure) is computed from the residual of the other terms, and represents the sink of anthropogenic CO_2 in natural land ecosystems. *From IPCC 2013, Figure 6.8, page 487, Ciais, P., C. Sabine, G. Bala, L. Bopp, V. Brovkin, J. Canadell, A. Chhabra, R. DeFries, J. Galloway, M. Heimann, C. Jones, C. Le Quere, R.B. Myneni, S. Piao and P. Thornton, 2013: Carbon and Other Biogeochemical Cycles, Figure 6.1. In:* ***Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*** *[Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA* (Ciais et al., 2013).

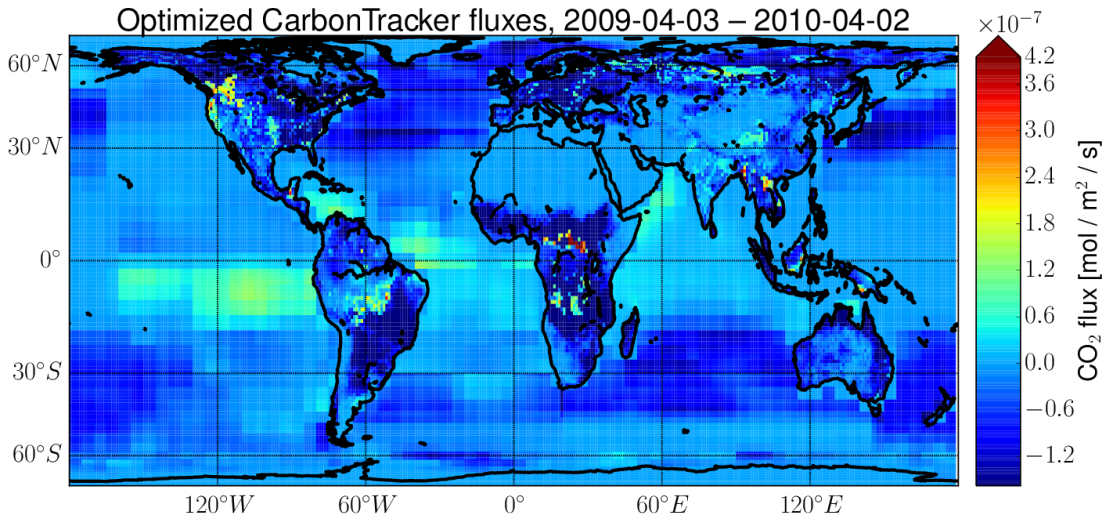


Figure 2.3: Biospheric and Oceanic Sources and Sinks of CO_2 from April 2009 to April 2010 as estimated by CarbonTracker. This high resolution map ($1° \times 1°$ latitude×longitude) is presented for illustrative purposes. Chapter 3 and 4 detail the uncertainties of carbon flux estimates.

As such, small relative errors in the strength of the source or sink add up to large errors in the net exchange. Also the exchange is hard to measure directly, because flux measurements are only representative for a small region and timespan (see section 2.3). Consequently our knowledge about sources and sinks of anthropogenic carbon is quite limited, often even calling in question whether a given region is a source or a sink.

To improve our understanding of net sources and sinks of carbon, we can use inverse models. Very roughly explained, they trace carbon dioxide on air currents and compare modelled CO_2 concentrations at measurement sites with actual measurements of atmospheric CO_2. Discrepancies between measurements and model are then attributed to sources and sinks of carbon upstream of the air currents (Peters et al., 2007). The resulting fluxes can yield results on fine spatial scales as shown in Figure 2.3, though as shown in chapters 3 and 4 uncertainties due to limited observation data require aggregation of the resulting fluxes on continental-scale regions to robustly distinguish net sources from net sinks.

Examples for further methods to improve the understanding of sources and sinks are discrimination between different plant types via their different isotope fractionation, as investigated for example by van der Velde et al. (2014), and assimilation of carbon monoxide measurements to discern biomass and fossil fuel burning from plant respiration (Gamnitzer et al., 2006).

However, these approaches show large differences, often larger than the actual estimated flux. For example the results from Keeling et al. (2011) suggest three times more uptake in the oceans than in the biosphere, while Peters et al. (2007) found that the biospheric sink is larger than the oceanic sink. These discrepancies make it necessary to get a better handle on actual flux uncertainties by using not only the results estimated using a single approach, but also uncertainties due to effects from different aspects of these approaches.

2.1.3 Prediction and Mitigation of Climate Change

The state of the art of the knowledge about uncertainties on climate change is available in the IPCC reports: The publications from the Intergovernmental Panel on Climate Change (Stocker et al., 2013). The IPCC reports are a collaborative work by scientists from all over the world and many different disciplines to collect and advance the current understanding about climate change and assess strategies to mitigate global warming and its adverse effects.

The reports aim at providing comprehensive information about the climate system in enough detail that political advisors can get a sufficiently complete understanding to judge how our current activities affect future climate and which measures can be taken to reduce unwanted consequences of our actions.

Expected consequences from current climate change include an increase of the global mean temperature and more and stronger tropical storms (IPCC, 2013), changes in rainfall patterns (Chadwick et al., 2013) and shifting of climatic regions (Kuang et al., 2014). In addition to the direct effects from the release of CO_2, increased air temperature could cause permafrost thaw on land (Khvorostyanov et al., 2008) and trigger a positive feedback which could stay active until 2300, releasing 600 to 1000 gigatons of carbon into the atmosphere – roughly 10 times the amount projected by the IPCC's high emission scenario (Schneider von Deimling et al., 2012). Seabed thawing could trigger the release of large amounts of methane (Portnov et al., 2013).

The political result from the research is that 20 years after the first IPCC report (IPCC Working Group I, 1996b) the governments of the largest carbon emitting countries committed to create a binding agreement to keep global warming below 2°C[3] to limit the probability of changes with high impact on economy and infrastructure. However, while this is an important step forward, Tschakert (2015) finish their assessment of the risks due to climate change with the conclusion "negative impacts of climate change under a 0.8°C temperature increase are already widespread, across the globe, and that danger, risk, and harm would be utterly unacceptable in a 2°C warmer world".

But despite broad scientific consensus about the information contained in the IPCC reports, these results have seen sustained attacks in the media and other non-scientific channels. Even the most fundamental result – the warming due to increased CO_2 concentrations in the atmosphere – is widely questioned.[4]

To improve communication between science and society, J. A. Curry (2011) argue for providing, among other dimensions, a likelihood estimate, representing "a probabilistic assessment of some well-defined outcome having occurred or occurring in the future."

This leads to the need for robust estimates of the accuracy of model results, uncertainties which take missing knowledge into account, to the need of providing actionable data by estimating the likelihood of specified scenarios from effects outside the knowledge of a single model, and the need to reduce the uncertainty of regional flux estimates. To this end, this thesis quantifies previously untested aspects of the uncertainties in flux estimation, and increases the amount of measurements which can be used together to reduce these uncertainties.

[3] The Guardian provides an overview of political action around climate change from the Kyoto protocol in 1997 which went in effect in 2004 over the targets created in Copenhagen 2009 to the Paris summit in 2015 where these plans are expected to become legally binding goals: `http://gu.com/p/49b3b/stw`

[4] For example the Guardian reported in the Article Propaganda trumps journalism in conservative media climate reporting (15 October 2015) about mischaracterization of an article by Ciuraru et al. (2015) in major newspapers as proof against global warning.

If the uncertainty of estimated fluxes can be sufficiently reduced, this should allow taking the biosphere into account for global regulation (like trading CO_2 certificates) to mitigate climate change.

2.2 Measuring greenhouse gases

The basic requirement for estimating sources and sinks of carbon from atmospheric concentrations are accurate concentration measurements of CO_2. For source-sink estimation there are several types of measurements. In-situ measurements, sampled with flasks or continuous methods, provide direct observations at a single location, while remote sensing from ground and remote sensing from satellite measure the total column of a trace gas in the atmosphere, either at a fixed location or – in the case of satellite measurements – at locations which vary with time.

This section describes the different measurement methods and their characteristics relevant to this work: in-situ measurements and their calibration against a common standard, ground based total column measurements, and total column measurements from satellites described.

2.2.1 In-situ measurements

The fundamental measurements – and the easiest to validate – are in-situ measurements. Air is either sampled in flasks and brought to a laboratory for analysis or directly pumped through a gas analyser.

The method employed throughout the "obspack" network used in this study (Masarie et al., 2014) is infra red absorption. The air is cooled to freeze out water, then a non-Dispersive Infra-red Analyser (NDIR) measures the absorption of infra-red light at 4.255 μm.[5]

Gas analysers are calibrated against reference air with known CO_2 concentration. This reference air is measured by freezing out CO_2 and N_2 from the air and measuring their temperature, expansion and pressure under heating to calculate the absolute concentration using physical expansion coefficients as described by Zhao et al. (1997). The uncertainty of the working standards used for this calibration is estimated at 0.071 μmol mol^{-1} with comparison against the five gravimetric standards from the National Institute for the Environmental Studies (NIES) in Tsukuba, Japan, showing an average and standard deviation of the differences of 0.004 $\pm$ 0.03 μmol mol^{-1} (Zhao and Tans, 2006).

In-situ measurements are conducted from many different platforms which define spatial and temporal coverage. The obspack distribution contains four main classes of platforms: weekly flasks at a fixed location, weekly flasks on ships, semi-continuous measurements (from several per day to several per hour) at a fixed location and semi-continuous measurements on aircrafts.

The aircraft measurements contain data from several specialized projects: The HIAPER Pole-to-Pole Observations (HIPPO, see Wofsy, 2011), Infrastructure for Measurements of the European Carbon Cycle (IMECC, see Feist et al., 2010; Messerschmidt et al., 2011), the Comprehensive Observation Network for Trace gases by Airliner (CONTRAIL, see Machida et al., 2008) and the Learjet flights described in Appendix A.3. Also included in the obspack distribution are measurements by instruments on small aircrafts from the National Oceanic and Atmospheric Administration/Earth System Research Laboratory (NOAA/ESRL) aircraft network.[6]

[5] As described by ESRL Global Monitoring Division - Global Greenhouse Gas Reference Network in Gas analysers (esrl.noaa.gov/gmd/ccgg/behind_the_scenes/meas_analyzers.html).

[6] Information about the NOAA/ESRL aircraft network is available from the website esrl.noaa.gov/gmd/ccgg/aircraft/

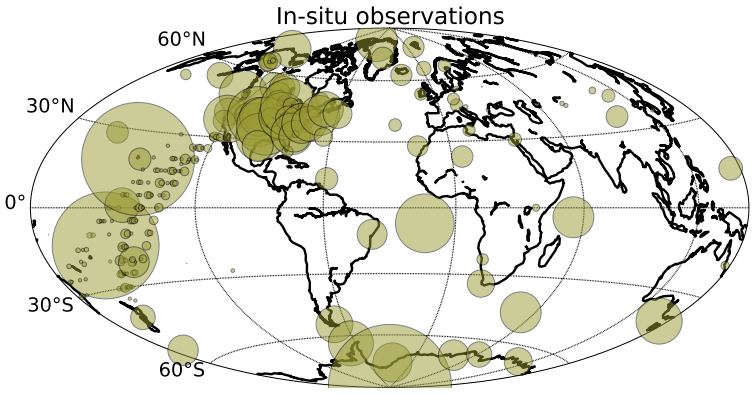


Figure 2.4: Distribution and data density of in-situ observation sites. The radius of a circle is proportional to $\frac{\sqrt{N}}{\sigma}$ with N the number of available measurements at the site in this location and σ their uncertainty estimate. The larger the symbol, the larger the weight of the site in the model.

At many but not all of the continuous measurement sites, the measurements are averaged to provide afternoon or night-time averages, using intra-day averaging periods representative of large scale fluxes and discarding single measurements outside the respective averaging periods. The averaging time period depends on the type of site, e.g., continental planetary boundary layer site or mountain site, and is chosen to minimize influence from complex local topology or meteorology.

The low uncertainty of the measurements is only valid for measuring the concentration at the point of the measurement. To get an uncertainty which is representative for the variability in a larger region, the uncertainty is increased to a fixed value for each site. This value accounts for measurement errors and for representativeness errors.

Representativeness errors originate from using the in-situ samples to represent the CO_2 concentration in box of 6° longitude and 4° latitude (roughly 500 km sidelength). Concentration uncertainties range from 0.75ppm for marine boundary layer sites over 2.5ppm for land sites up to 7.5ppm for sites which experience variable meteorological conditions like varying wind from sea and from land. The study in chapter 3 excludes some of these sites for additional validation. Table A.1 lists the in-situ observation records used in chapter 3. Figure 2.4 shows the global distribution of in-situ observations used in chapter 4 together with a visual representation of their weight due to sampling frequency and representativeness error.

2.2.2 Remote sensing from ground

While CO_2 concentration in the atmosphere is regularly sampled via the well established surface air sampling network (Gurney et al., 2002), sparseness of this observation network limits the representativeness of the measured mixing ratios for large scale variation of CO_2. In addition, the in-situ surface sampling network only provides a limited view of the atmosphere, since only few sites are located above the boundary layer and none reach the middle of the troposphere. Aircraft campaigns provide a way to test how well the model captures the CO_2 concentration in the middle to upper troposphere, but they are even sparser in time and space than the surface sampling sites.

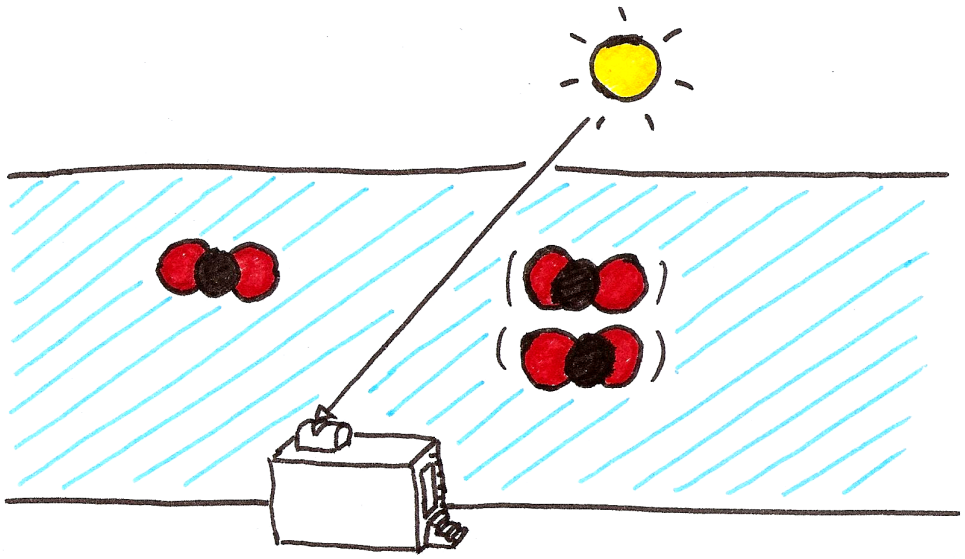


Figure 2.5: Schematic for direct sunlight remote sensing from the ground.

Remote sensing measurements from ground provide a part of this missing information. They differ from in-situ absorption spectrometry by looking into a remote light source like the sun as shown in Figure 2.5 and calculating the absorption by fitting several parameters to a spectral measurement. The most widespread system for accurate remote sensing of CO_2 is the Total Carbon Column Network (TCCON). It employs high resolution Fourier Transform Spectrometry (FTS) and all its sites evaluate near infra-red (NIR) measurements with the same software package (GGG from Toon et al., 2009b).

Other Fourier Transform Infra-red (FTIR) networks include the Network for the Detection of Atmospheric Composition Change (NDACC) which operates in the mid infra-red (MIR) and uses different instruments and evaluation methods. Its heterogeneity – using different instruments and retrieval methods – makes its measurements harder to assimilate in a global model, but it includes sites which already provided observations in 1991 and which measure a large number of species. The NDACC network does not provide public CO_2 data yet, but Barthlott et al. (2015) analysed the quality of CO_2 retrieved from sites which coincide with TCCON measurements, finding good stability over the three years of the comparison, though with a bias of 2.5%. Finally IMK-ASF at KIT is establishing the Collaborative Carbon Column Observation Network (CoCCON) built from small FTIR spectrometers which are lower spectral resolution (and an order of magnitude cheaper) than those used in the TCCON network. Some of these instruments were already used in studies to detect greenhouse emissions of the city Berlin, Germany (Hase et al., 2015).

More information about FTS is available in Fourier Transform Spectroscopy by Davis et al. (2001) and from Griffith and de Haseth (2007).

When comparing remote sensing retrievals to modelled concentration profiles, the vertical sensitivity of the measurement needs to be taken into account. This sensitivity is generally called the averaging kernel ($\mathbf{A}$), defined via

$$\hat{\vec{x}} = \mathbf{A}\vec{x}_{\text{true}} + (\mathbf{1} - \mathbf{A})\vec{x}_b \tag{2.1}$$

with the retrieval $\hat{\vec{x}}$, the averaging kernel $\mathbf{A}$, the true profile $\vec{x}_{\text{true}}$ and a prior $\vec{x}_b$. For total columns, $\vec{x}_{\text{true}}$ reduces to a scalar and $\mathbf{A}$ reduces to a vector $\vec{A}_c$ as calculated by

$$\vec{A}_c = \vec{h}^T\mathbf{A} \tag{2.2}$$

with the pressure weighting function $\vec{h}$. The results estimated by ground-based remote sensing sites are validated against aircraft campaigns which measure the carbon profiles with in-situ instruments. The TCCON network reports an accuracy and precision better than 0.25% for the aggregated total column (Wunch et al., 2011a). Representativeness errors for the TCCON network are estimated in section 4.3.

Ground based remote sensing networks provide measurements at a high temporal resolution (several measurements per hour), but due to the cost of the instruments only at a limited number of fixed positions worldwide. In 2014 the TCCON network had 18 operational sites. Due to measuring the total column, they are less affected by local effects like biospheric respiration in close vicinity to the site than in-situ measurements.

2.2.3 Remote Sensing from satellite

Satellite-based remote sensing column measurements work similar to ground based measurements. As shown in Figure 2.6 their light source is not the direct sunlight, but the reflected sunlight from the Earth's surface. As such they have to take into account effects like multiple scattering.

Current remote sensing instruments for CO_2 include TANSO[7] on board of the GOSAT sattelite, launched in 2009 and still in extended operation and the spectrometer of the OCO-2 satellite, launched in 2014. Software systems for calculating columns of CO_2 and other trace gases from GOSAT satellites include RemoTeC (Butz et al., 2011; Butz et al., 2013), ACOS-GOSAT (O'Dell et al., 2012), BESD (Reuter et al., 2011), the algorithm from the National Institute for Environmental Studies, Japan (Yoshida et al., 2011, 2013), and from Parker et al. (2011).

Column retrievals from GOSAT and OCO-2 are validated against TCCON measurements. The results from the RemoTeC algorithm exhibit a scatter of 2.5 to 2.8 ppm against collocated TCCON observations with a station-to-station bias of 0.84 (Dils et al., 2014; Butz et al., 2011).

Satellite measurements from GOSAT provide high spatial resolution, with tens of thousands of measurement locations worldwide, but varying with cloud cover and with limited temporal resolution – nominally returning to each position after 3 days, but practically much less due to interference from clouds. As such they complement ground based total column measurements which have limited spatial coverage and high temporal resolution.

2.2.4 Summary

The available measurements differ in their measurement geometry (in-situ or remote sensing), their sensitivity (point measurement or total column), their coverage (fixed position or moving platform) and their frequency (from once per week to several per hour). Their characteristics are summarized in table 2.1.

[7] The name TANSO means carbon in Japanese, shown in Figure 2.6.

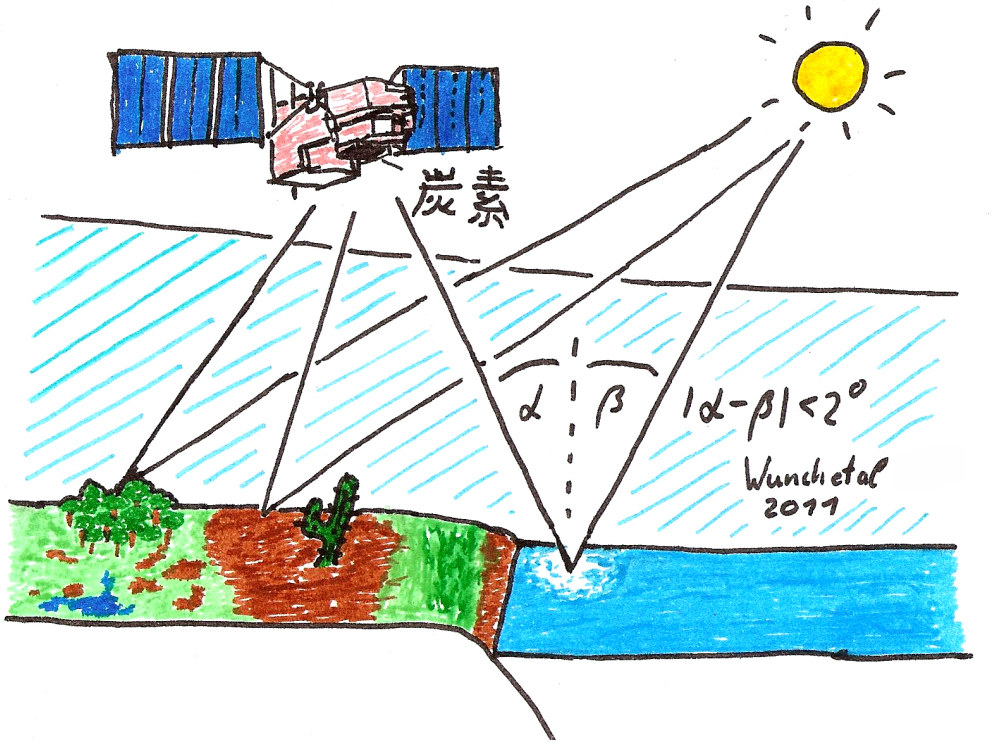


Figure 2.6: Remote sensing from satellite using undirected reflection at the ground and glint reflection on the ocean. The point where the satellite sees reflection of the sun on the approximately planar ocean is called glint spot. It is defined by Wunch et al. (2011a) as the sea surface area which has a maximum difference between the angles of incoming and outgoing light of 2°.

The classes of measurements build on each other as validation. Satellites are validated by ground based column measurements which in turn get checked against in-situ aircraft measurements. All in-situ measurements are calibrated against a common standard air which is measured in absolute units via manometric methods.

2.3 Bottom-up modelling of CO_2 surface fluxes

Despite the high quality of atmospheric concentration measurements, the available spatial and temporal data coverage is still far too limited to directly constrain the global fluxes. To fill the gaps in the data and to keep the modelled fluxes within plausible limits, global flux estimation builds upon bottom-up models. These use direct measurements of the surface flux of CO_2, but suffer from of limited representativeness of the flux measurements and high variability of the net fluxes as visualized in Figure 2.2.

2.3.1 Biosphere

Measurement methods for biosphere fluxes include – ordered by increasing spatial coverage – leaf cuvettes, plant chambers, soil chambers and eddy covariance measurements. For leaf cuvettes, plant chambers and soil chambers, a leaf, plant or soil region is placed in an enclosed space

to directly measure the change in trace gas concentrations due to biospheric activity (see Fig. 2.7a). They require significant experimental investment since the area they measure is very small but they are straightforward to evaluate. For eddy covariance measurements, the trace gas concentration and wind velocity are measured at several height levels to estimate the flux from covariances between the concentrations and wind velocities (see Fig. 2.7b). Eddy covariance measurements are representative of a region spanning from a few hundred meter to several kilometre upwind (Baldocchi, 2003). Evaluation of eddy covariance methods uses the assumption of a flat surface (Curtis et al., 2002; Papale et al., 2006; Göckede et al., 2008).

Direct flux measurements are coordinated by several groups around the world. The FLUXNET project[8] collects and provides data from over 450 active measurement sites worldwide (for details see the readme document for Luyssaert et al., 2009).[9] Biometric methods to estimate net fluxes include litter fall to estimate the change in stored carbon (Davidson et al., 2002; Curtis et al., 2002).

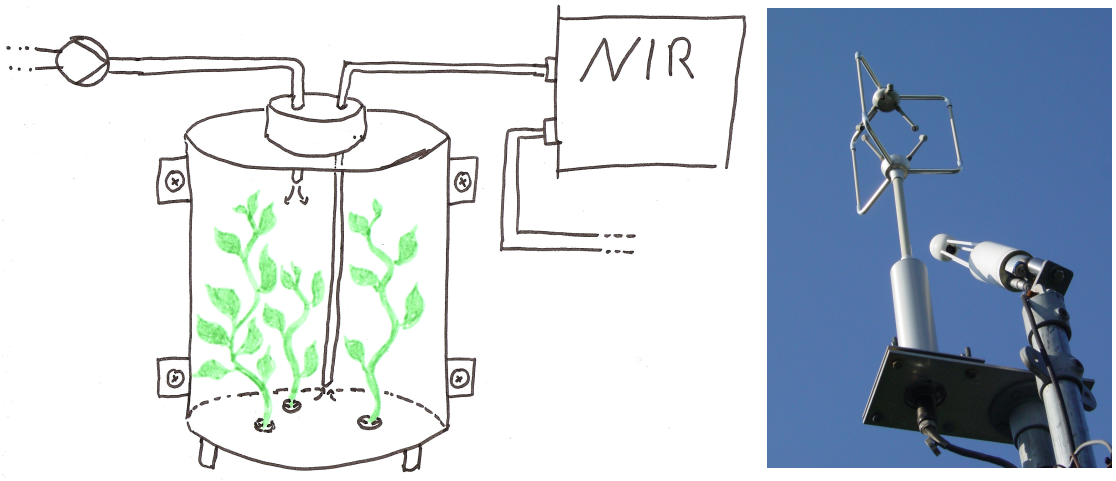
Due to the limited footprint of the flux measurement methods, they can only get data representative of a limited region, so they require up-scaling to global scales. This study employs the Simple Biosphere model using the Carnegie-Ames-Stanford Approach (SIBCASA as by Schaefer et al., 2008), described by Peters (2014): "SIBCASA is a carbon cycle model that represents the uptake of CO_2 by different types of vegetation and its subsequent transfer back to the atmosphere through autotrophic and heterotrophic respiration. Its mechanistic description of the processes involved is driven by a combination of high-resolution weather data and satellite remote sensing products and includes interactions between the carbon, water, and energy cycles of the land-surface". The resulting CO_2 concentrations when using SIBCASA fluxes, however, do not match the concentrations exactly. On an 18 months run, atmospheric CO_2 concentrations which start at a state consistent with the measurements at the background station at Mauna Loa, Hawaii, overestimate the real measurements at the end of the run by more than 1 ppm.

[8] The data from the FLUXNET project is published by the Distributed Active Archive Center For Biogeochemial Dynamics (ORNL DAAC) and available on daac.ornl.gov/FLUXNET/fluxnet.shtml.

[9] Currently 468 eddy covariance towers organized in the FLUXNET project are active, as reported at fluxnet.ornl.gov/site_status.

Table 2.1: summary of the discussed measurements with precision describing the statistical scatter around the measured value and accuracy describing the estimated systematic error.

technique	precision	accuracy	sensitivity	additional information
in-situ ground	0.071 ppm	0.03 ppm	points on ground	obspack (Masarie et al., 2014; Zhao and Tans, 2006)
in-situ aircraft	0.071 ppm	0.03 ppm	points in vertical profiles	HIPPO (Wofsy, 2011), CONTRAIL (Machida et al., 2008), IMECC (Feist et al., 2010; Messerschmidt et al., 2011), Learjet (see A.3)
remote-sensing ground	<1 ppm	<1 ppm	total columns	TCCON (Toon et al., 2009a), CoCCON (Hase et al., 2015)
remote-sensing satellite	2.5 ppm	0.84 ppm	total columns	GOSAT (Dils et al., 2014; Butz et al., 2011; Reuter et al., 2011; O'Dell et al., 2012; Yoshida et al., 2011, 2013)



(a) (b)

Figure 2.7: Measuring fluxes directly. Fig. 2.7a shows the schematic representation of a leaf bag. Air is pumped through the bag and the CO_2 concentrations are measured with a near infra-red spectrometer. Fig. 2.7b shows an open path eddy covariance system consisting of an ultrasonic anemometer to measure 3D air flow and an infra-red gas analyser (IRGA) to measure carbon dioxide (Photo by Veedar, in the public domain).

Alexandrov (2014) analysed the representativeness of flux sites with a simplified system and found that different regions on the globe need different parameters to fit atmospheric CO_2 measurements, highlighting the need to use atmospheric measurements to validate and calibrate flux models – and to test whether climatic changes affect the reaction of the biosphere, requiring new adjustments to the models.

2.3.2 Oceans

Ocean fluxes are modelled with a combination of ocean circulation models and atmospheric transport.

The ocean can store large amounts of CO_2 due to acid buffering. When CO_2 is taken up by the ocean's surface, the carbon is stored in three different acid compounds, as shown in equations (2.3), (2.4) and (2.5): carbonic acid (H_2CO_3), the bicarbonate ion (HCO_3^-), and the carbonate ion (CO_3^{2-}). This buffer adjusts to the atmospheric concentrations of CO_2. Measuring the acidity of the ocean can act as a proxy for dissolved inorganic carbon (DIC). Due to different transport velocities in the ocean, increased CO_2 concentrations have different regional penetration into the deeper ocean, ranging from a few hundred meters in tropical latitudes to one kilometre in northern pacific and the southern Atlantic (Jacobson et al., 2007). The equilibrium between DIC and atmospheric CO_2 depends on the water temperature, with warming water releasing CO_2 into the atmosphere and cooling water taking up atmospheric CO_2. As such the ocean circulation and temperature affect atmospheric concentrations of CO_2.

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \tag{2.3}$$

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^- \tag{2.4}$$

$$HCO_3^- \rightleftharpoons H^+ + CO_3^{2-} \tag{2.5}$$

This study uses the Ocean Inversion Fluxes (oif) from Jacobson et al. (2007). They combine the Transcom3, Level 1 transport (Gurney et al., 2002) with an ocean inversion and use atmospheric measurements from the GLOBALVIEW dataset (NOAA Global Monitoring Division: Boulder, Colorado, U.S.A., 2013) and ocean measurements from the GLODAP project (Sabine et al., 2005) which aggregates and harmonizes data from hydro-graphic stations active from 1985 to 1999 and from 1972 to 1990, collected in the World Ocean Circulation Experiment (WOCE Data Products Committee, 2002), Joint Global Ocean Flux Study (Doney et al., 2002) and NOAA Ocean-Atmosphere Exchange Study (OACES).

The model first uses concentrations with removed anthropogenic part to estimate steady state ocean transport and then transports the anthropogenic concentration enhancements to discriminate between air-sea fluxes and internal ocean fluxes. This assumes that the uptake of anthropogenic CO_2 increases proportional to the mismatch between atmospheric and oceanic CO_2 partial pressure.

2.3.3 Fire and Fossil Fuel Fluxes

In addition to the biospheric and oceanic fluxes, this study also uses fluxes representing fires and fossil fuel burning. Aggregated fluxes for one year are shown in figures 2.8 and 2.9.

Fire fluxes are derived from the Global Fire Emissions Database (GFEDv2, van der Werf et al., 2010), while fossil fuel burning fluxes use the Miller dataset (Peters et al., 2007, and its supplement).

The different kinds of fluxes are summarized in table 2.2.

2.4 Atmospheric Transport of CO_2

To adjust fluxes to fit atmospheric concentration measurements, the fluxes need to be connected to the concentration measurements with a tracer transport model. The core of these models is a representation of meteorology, either as dynamic system or using pre-calculated mass transport data – wind fields.

Table 2.2: A priori fluxes used in the inverse models.

flux type	model	reference
biosphere	SiBCASA	Schaefer et al. (2008)
ocean	OIF	Jacobson et al. (2007)
fire	GFED	van der Werf et al. (2010)
fossil	Miller	Peters et al. (2007) and its supplement

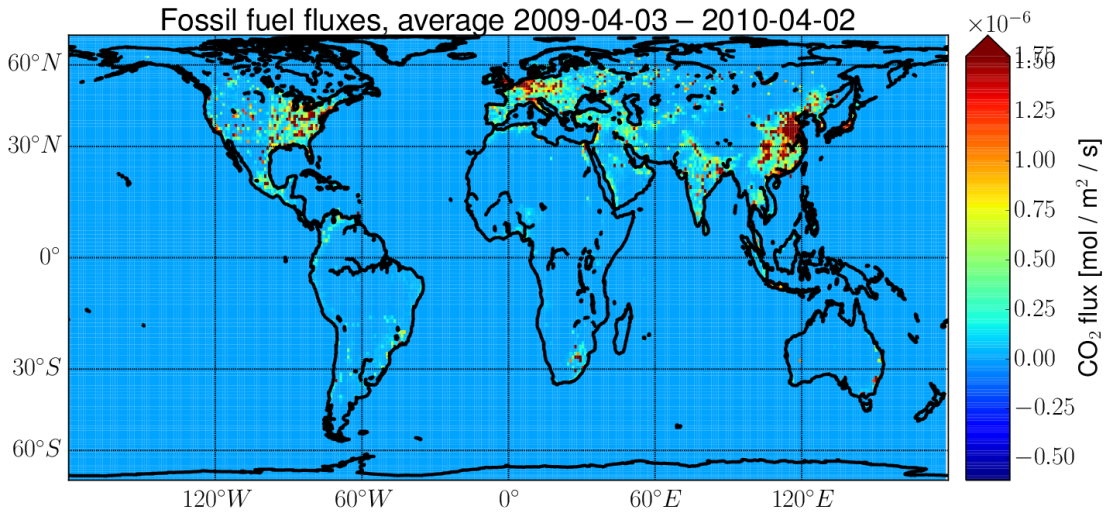


Figure 2.8: 1°×1° map of Fossil Fuel fluxes for CarbonTracker from April 2009 to April 2010.

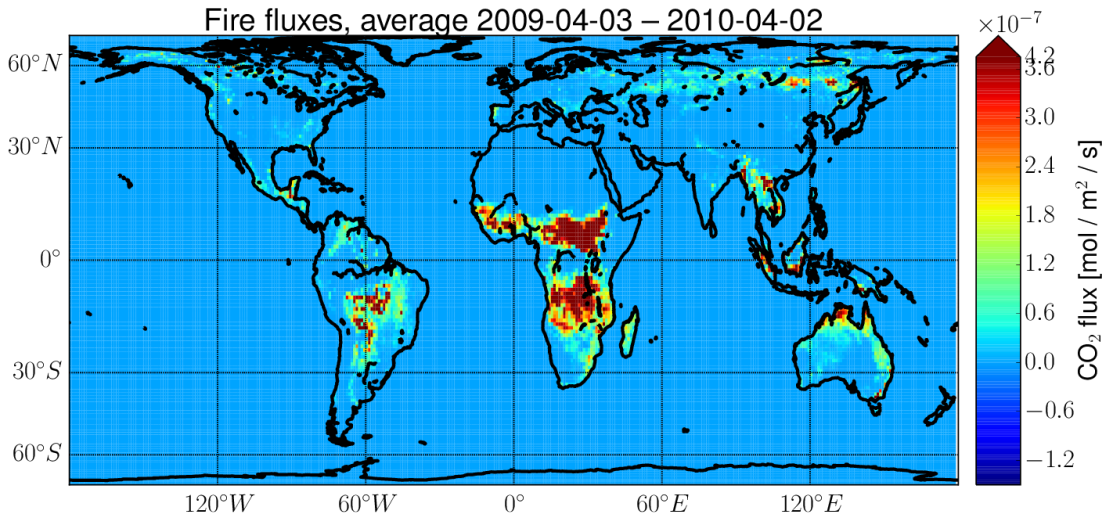


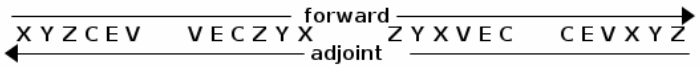
Figure 2.9: 1°×1° map of Fire fluxes for CarbonTracker from April 2009 to April 2010.

2.4.1 Meteorological data

Tracer transport models use meteorological data to transport chemical tracers, allowing for comparison of emissions and chemistry with atmospheric measurements.

This study uses meteorological data of the Integrated Forecast System (IFS) from the European Centre for Medium-Range Weather Forecasts (ECMWF) with the ERA Interim reanalysis (Dee et al., 2011) using a horizontal resolution of 80 km grid spacing and 60 vertical pressure levels. To model mass flow, the IFS model uses different approximations for the different directions. In vertical dimension it employs a finite elements method while the horizontal flow uses spectral

Figure 2.10: Schematic representation of the operations in the tracer model 5 (TM5), with the operations X,Y,Z-advection, chemistry (C), emission (E) and vertical transport (V: convection and diffusion, parametrised). for details see the Supplement of Meirink et al. (2008).



discretization, representing modes of the atmosphere as Legendre functions on a reduced Gaussian grid (Hortal and Simmons, 1991).

2.4.2 Transporting tracers

The mass flow from the meteorological data allows transporting tracers to estimate their evolution with time. In this study, the Transport Model 5 (TM5, Krol et al., 2005) is used. It is a Eulerian chemical transport model with optional two-level nesting of higher resolution regions in a lower resolution global grid. TM5 splits the tracer transport into X,Y,Z-advection, chemistry (C), emission (E) and vertical transport (V: convection and diffusion), chaining the operations as shown in figure 2.10 (Meirink et al., 2008). It provides an adjoint model which yields sensitivity information of concentrations to emissions. This fulfils the requirement for implementing inverse modelling via four dimensional variation methods (4DVar, see section 2.5.4). To minimize discretization errors and oscillating behaviour, TM5 uses the slope between the concentration of neighbouring grid boxes with a spatial leapfrog scheme as base for its transport operators (Russell and Lerner, 1981).

As shown by Patra et al. (2008), TM5 is among the models with the highest correlation of modelled CO_2 concentrations and measurements on hemispheric scale, making it a sane choice for this comparison.

Current developments in high resolution atmospheric tracer transport (see, for example, Sofiev et al., 2015) might allow for a reduction of the representativeness errors ascribed to measurements (described for in-situ measurements in section 2.2.1 and for TCCON in section 4.3). However, as shown in chapter 3, this is not guaranteed to yield better flux estimates, because the main source of uncertainty in current estimates is the low global coverage of observational data, giving too little information to derive robust fluxes on regional scales.

2.5 Assimilation Methods

Data assimilation (DA) combines different sources of measurements to estimate a state which is most consistent with measurements and prior knowledge. The data assimilation methods in this thesis estimate the surface fluxes of CO_2 using a transport model which calculates atmospheric concentrations from these fluxes.

2.5.1 Different Approaches: ensemble or variational

There are two main classes of data assimilation techniques for complex inversions, variational methods and ensemble methods (Lahoz et al., 2007; Lahoz and Schneider, 2014). Chapter 3 of this study compares an ensemble method and a variational method to select one of them for going forward.

Both approaches are approximate variants of the general Bayesian optimal estimation scheme (e.g. Rodgers, 2000) which aims at balancing prior or background information with actual measurement information to derive robust parameter estimates. But they make different trade-offs which can yield different deviations from the optimal solution dependent on the background uncertainty, the observation density and the required computational efficiency (Fairbairn et al., 2013).

Approximations are necessary to render the inverse problem computationally feasible since real-world CO_2 surface flux inversions typically involve thousands of concentration measurements and millions of unknown flux parameters. Both schemes can either treat the entire considered assimilation period at once or divide it into shorter periods to be treated sequentially. Ensemble methods approximate the exact solution from an ensemble of model runs, while variational methods approach the optimal solution step-by-step (e.g. Juhász and Bölöni, 2007; Gilbert and Lemaréchal, 1989).

The performance of ensemble methods and variational methods has been evaluated previously for numerical weather prediction (e.g. Kalnay, 2005) and direct optimization of atmospheric gas abundances (Skachko et al., 2014). Chatterjee and Michalak (2013) are the first to evaluate the performance of the two methods for the purpose of CO_2 surface flux estimation. However, they use a synthetic setup with simulated observations and a 1-dimensional transport model which has the advantage of knowing the true fluxes and for which a direct Bayesian inversion is computationally feasible, but which does not capture the complexities of real-world flux estimation. In particular they find that under constraints on model runtime and resource use, the estimated surface fluxes are more realistic with their variational implementation than with their ensemble method, and that for both models small-scale fluxes (flux aggregation spanning up to 5% of the model size) are very sensitive to the data coverage and distribution.

The models used in this study are TM5-4DVar using a variational method which treats the whole assimilation period at once, and CarbonTracker, an ensemble method which steps through the assimilation period in overlapping assimilation time windows.

The approach by TM5-4DVar allows for much longer temporal correlation than the approach used by CarbonTracker, because it includes explicit decaying temporal correlation whereas CarbonTracker relies on the definition of an assimilation window with implicit temporal correlation (see section 2.5.3). However the variational approach has more complicated requirements on the transport model (it needs an adjoint model). The difference between ensemble and variational algorithms when applied to inverse flux modelling using real measurements has not been tested prior to this study.

From the results of Chatterjee and Michalak (2013), we expect differences in attribution of fluxes near sources as well as for fluxes which are influenced by long-term effects like inter-hemispheric transport.

2.5.2 Assimilating measurements

Inverse modelling uses CO_2 concentration gradients observed in the Earth's atmosphere to quantify the spatio-temporal distribution of the net CO_2 surface fluxes (e.g. Enting, 2000; Peters et al., 2007; Chevallier et al., 2010; Feng et al., 2011; Peylin et al., 2013). To this end, various data assimilation (DA) techniques have been developed.

Data assimilation systems (DA systems) aim at inferring a state vector $\vec{x}$ that contains the value to be optimized. In the case of surface flux data assimilation these are spatially and temporally binned surface fluxes or a related quantity such as scaling factors for an initial guess flux field. To this end, the systems exploit measurements chained into an observation vector $\vec{y}$. For surface

flux optimization, these are observations of the atmospheric concentration. The state and the observations are linked through the transport and observation operator $\mathbf{H}$ which is linear for the case of our CO_2 flux inversions, but in general could be non-linear such as for CH_4 flux inversions. The inverse method reconstructs an optimized state vector $\hat{\vec{x}}$ (along with uncertainty estimates) from the measurements and the transport and observation operator.

Typically, the inverse problem of estimating $\hat{\vec{x}}$ from a set of observations $\vec{y}$ is ill posed. Due to sparse observational coverage, measurement errors or measurement configuration, the observations contain insufficient information to determine all components of $\hat{\vec{x}}$ independently. A background flux estimate $\vec{x}_\mathrm{b}$ from biosphere and ocean models is used to provide a constraint that fills the null-space where measurement information is insufficient. Accordingly, the state vector of fluxes $\vec{x}$ is determined by minimizing a cost function J that typically consists of two terms, the mismatch between measured and modelled observations and the mismatch between the fluxes to be estimated and the background estimate,

$$\begin{aligned} J = \quad & (\vec{y} - \mathbf{H}\vec{x})^T \mathbf{R}^{-1} (\vec{y} - \mathbf{H}\vec{x}) + \\ & + (\vec{x} - \vec{x}_\mathrm{b})^T \mathbf{B}^{-1} (\vec{x} - \vec{x}_\mathrm{b}) \end{aligned} \tag{2.6}$$

with $\mathbf{R}$ the observation covariance and $\mathbf{B}$ the background flux covariance. $\mathbf{R}$ and $\mathbf{B}$ define the relative weights of the measurement and background mismatch.

In general, minimization of Eq. (2.6) can be solved by means of matrix algebra (Rodgers, 2000), yielding optimized fluxes and their error covariances,

$$\hat{\vec{x}} = \vec{x}_\mathrm{b} + \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (\vec{y} - \mathbf{H}\vec{x}_\mathrm{b}) \tag{2.7}$$

$$= \vec{x}_\mathrm{b} + (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\vec{y} - \mathbf{H}\vec{x}_\mathrm{b}), \tag{2.8}$$

$$\hat{\mathbf{B}} = \mathbf{B} - \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \mathbf{H}\mathbf{B} \tag{2.9}$$

$$= (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1})^{-1}, \tag{2.10}$$

with $\hat{\vec{x}}$ the a posteriori state vector and $\hat{\mathbf{B}}$ the respective covariance matrix. Equivalence of equations (2.7) and (2.8) can be shown following Eqs. (4.11), (4.12) and (2.27) from Rodgers (2000). Using

$$\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) = (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{B}\mathbf{H}^T \tag{2.11}$$

(Rodgers, 2000, Eq. 4.11) equivalence is shown by multiplying from left with $(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}$ and multiplying from right with $(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$:

$$(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}. \tag{2.12}$$

This minimization can be direct or indirect. Direct assimilation uses a state which corresponds to the measurements such as the concentration of CO_2. Here the model fills the null-space where measurements are missing. This method is typically used in numerical weather prediction with current research going towards heterogeneous high resolution assimilation like dispersion of birch pollen (Sofiev et al., 2015).

Indirect assimilation uses a state from which the measured quantity is derived, like sources and sinks of carbon. In this study the result corresponds to the first order derivative, the change in the transported tracer. For tracers like CO_2 which have a high absolute concentration compared to the changes due to sources and sinks, a small change in the tracer can lead to strong changes in the regional attribution of sources and sinks, and missing information can lead to incorrect attribution.

While theoretically minimizing Eq. (2.6) reduces to a matrix inversion for linear systems like CO_2 flux inversion (e.g. Rodgers, 2000), the large number of parameters to be estimated and the amount of measurements to be ingested require approximate methods which are numerically efficient such as the Ensemble Square Root Filter (EnSRF) and four dimensional variation (4DVar).

2.5.3 Ensemble Method: CarbonTracker

CarbonTracker is an inverse modelling framework based on the Ensemble Square Root Filter (Whitaker and Hamill, 2002) developed by Peters et al. (2007). Instead of solving the minimization problem in one step, the EnSRF determines optimized surface fluxes sequentially in a time stepping approach with $\vec{x}_t$ defining a subset of $\vec{x}$ for a certain time window. In the setup for chapter 3, $\vec{x}$ contains scaling factors for the surface fluxes for 96 weeks, while $\vec{x}_t$ only spans 5 weeks.

Commonly, a gain matrix $\mathbf{G}$ is defined as

$$\mathbf{G} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \tag{2.13}$$

$$= (\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1})^{-1}\mathbf{H}^T\mathbf{R}^{-1}. \tag{2.14}$$

Equations (2.7) and (2.9) then read

$$\hat{\vec{x}}_t = \vec{x}_{\mathrm{b},t} + \mathbf{G}_t\left(\vec{y}_t - \mathbf{H}_t\vec{x}_{\mathrm{b},t}\right), \tag{2.15}$$

$$\hat{\mathbf{B}}_t = \mathbf{B}_t - \mathbf{G}_t\mathbf{H}_t\mathbf{B}_t \tag{2.16}$$

with the partial Gain Matrix

$$\mathbf{G}_t = \mathbf{B}_t\mathbf{H}_t^T\left(\mathbf{H}_t\mathbf{B}_t\mathbf{H}_t^T + \mathbf{R}_t\right)^{-1} \tag{2.17}$$

where subscript t indicates quantities of reduced dimensions, for the time step under investigation.

In layman's terms the gain matrix transforms a mismatch between the model estimate $\mathbf{H}\vec{x}_{\mathrm{b},t}$ and observations $\vec{y}$ into the most probable adjustment of the flux scaling factors $\vec{x}_{\mathrm{b},t}$, taking the uncertainty of the observations $\mathbf{R}$ and the uncertainty of the model $\mathbf{B}$ into account.

Once Eqs. (2.15) and (2.16) are solved for time slice t, the solution of the scaling factors $\hat{\vec{x}}_t$ is used as the background estimate $\vec{x}_{\mathrm{b},t+1}$ for the next time slice $t+1$, assuming that a simple persistence forecast is adequate for our CO_2 flux inversion problem as shown in Figure 2.11. At each optimization step the oldest cycle at the "end" of the state vector drops out of the state vector and is used as a posteriori flux estimate while a new cycle is added to the "beginning" of the state vector (see Figure 2.11). As such, each one-week cycle experiences a number of optimization steps equal to the number of weeks in the assimilation time window (Peters et al., 2005, section 2.3).

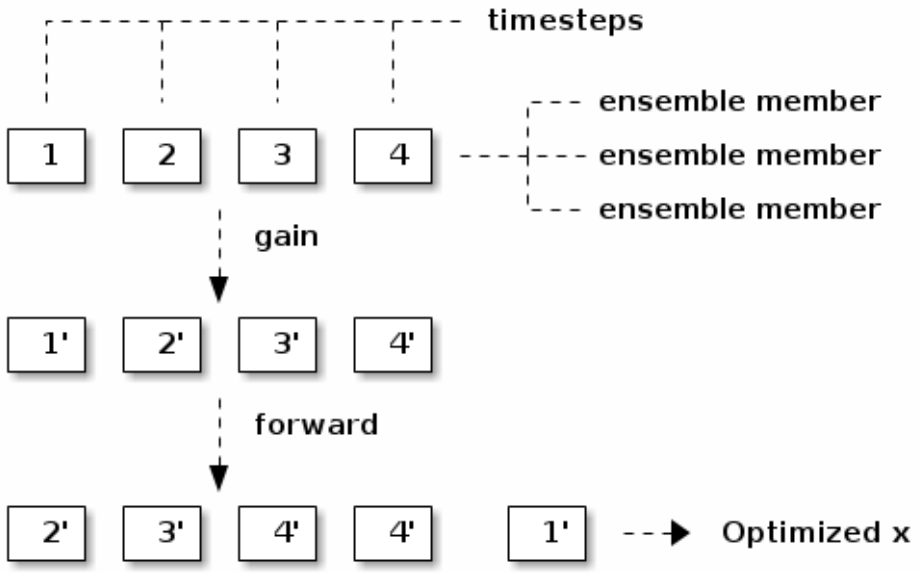


Figure 2.11: Simplified visualization of the time stepping approach employed by CarbonTracker. For each time step the state is represented by a number of ensemble members whose mean value and spread are consistent with the mean value and uncertainty of the surface flux scaling factors. In the gain step, CarbonTracker adjusts the ensemble members to fit atmospheric measurements. In the forward step, it provides the oldest time step as result (here 1') and adds the scaling factors for the fluxes in the next time step. The number of time steps in CarbonTracker is defined by the lag (default: 5), their length by the cycle fluxes (default: one week).

The covariance $\mathbf{B}_{t+1}$ is prescribed at each time step as described in Peters et al. (2005). Given an initial guess for the first background state, this strategy allows for sequentially calculating the complete state vector $\hat{\vec{x}}$.

To estimate the gain matrix $\mathbf{G}_t$, the EnSRF uses an ensemble approach. The ensemble members $\vec{x}^i_{\mathrm{b},t} = \vec{x}_{\mathrm{b},t} + \Delta\vec{x}^i_{\mathrm{b},t}$ ($i = 1 \ldots E$, with E the ensemble size) of the background state are drawn such that their mean and covariance is consistent with the background state $\vec{x}_{\mathrm{b},t}$ and background covariance matrix $\mathbf{B}_t$, respectively, so that

$$\begin{aligned} \mathbf{B}_t \approx \; & \frac{1}{E-1}\left(\Delta\vec{x}^1_{\mathrm{b},t}, \Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \Delta\vec{x}^E_{\mathrm{b},t}\right) \\ & \cdot \left(\Delta\vec{x}^1_{\mathrm{b},t}, \Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \Delta\vec{x}^E_{\mathrm{b},t}\right)^T \end{aligned} \tag{2.18}$$

,

with each of the vectors $\left(\Delta\vec{x}^1_{\mathrm{b},t}, \Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \Delta\vec{x}^E_{\mathrm{b},t}\right)$ defining the deviations from the mean state. In CarbonTracker the required calculation of the "square root" of $\mathbf{B}$ is realized with a Cholesky decomposition (Cholesky, 1910).

Then, the terms $\mathbf{H}_t\mathbf{B}_t\mathbf{H}_t^T$ and $\mathbf{B}_t\mathbf{H}_t^T$ required for calculating $\mathbf{G}_t$ following Eq. (2.17) are approximated using the results from an ensemble run of the possibly non-linearized transport model $\mathcal{H}$

$$\begin{aligned}\mathbf{H}_t\mathbf{B}_t\mathbf{H}_t^T \approx\ & \frac{1}{E-1}\left(\mathcal{H}_t\Delta\vec{x}^1_{\mathrm{b},t}, \mathcal{H}_t\Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \mathcal{H}_t\Delta\vec{x}^E_{\mathrm{b},t}\right) \\ & \cdot \left(\mathcal{H}_t\Delta\vec{x}^1_{\mathrm{b},t}, \mathcal{H}_t\Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \mathcal{H}_t\Delta\vec{x}^E_{\mathrm{b},t}\right)^T \qquad (2.19) \\ \mathbf{B}_t\mathbf{H}_t^T \approx\ & \frac{1}{E-1}\left(\Delta\vec{x}^1_{\mathrm{b},t}, \Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \Delta\vec{x}^E_{\mathrm{b},t}\right) \\ & \cdot \left(\mathcal{H}_t\Delta\vec{x}^1_{\mathrm{b},t}, \mathcal{H}_t\Delta\vec{x}^2_{\mathrm{b},t}, \ldots, \mathcal{H}_t\Delta\vec{x}^2_{\mathrm{b},t}\right)^T, \qquad (2.20)\end{aligned}$$

where the approximation becomes more exact with increasing ensemble size E.

The EnSRF method can yield robust results with non-linear transport operators $\mathcal{H}$ as long as the transport model is close to linear for small perturbations ($\mathcal{H}(\vec{x}+\vec{\Delta x}) \approx \mathbf{H}\vec{x} + \mathbf{H}\vec{\Delta x}$). Using Eqs. (2.19) and (2.20), the gain matrix $\mathbf{G}_t$ can be calculated from Eq. (2.17) to update the mean state estimate $\hat{\vec{x}}_t$ via Eq. (2.15).

In addition to the mean state, the ensemble deviations are updated via

$$\Delta\hat{\vec{x}}_{\mathrm{b},t} = \Delta\vec{x}_{\mathrm{b},t} - \alpha\mathbf{G}\cdot\mathbf{H}(\Delta\vec{x}_{\mathrm{b},t}) \qquad (2.21)$$

with the scaling factor α calculated from the model uncertainty $\mathbf{HBH}^T$ and the measurement uncertainty following

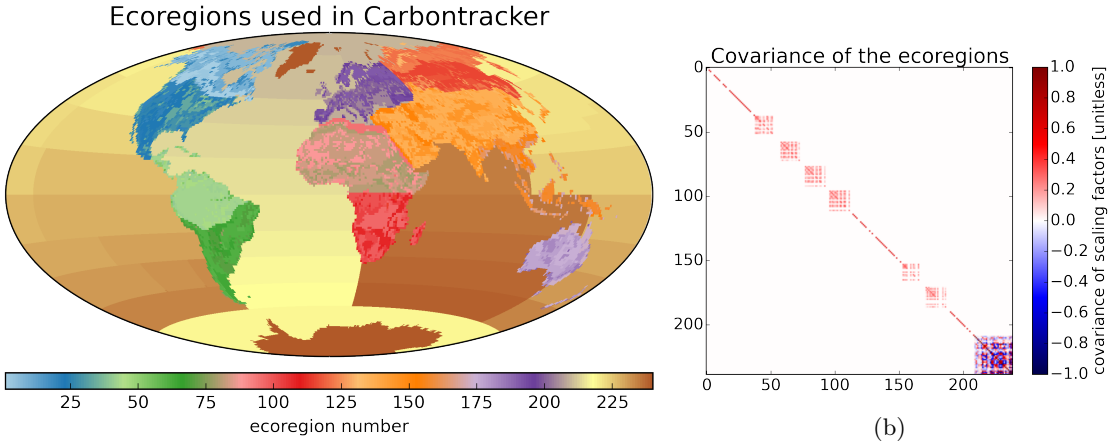
$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{HBH}^T + \mathbf{R}}}\right)^{-1}. \qquad (2.22)$$

Peters et al. (2005) and Whitaker and Hamill (2002) describe the derivation of α from the requirement that the optimized covariance $\hat{\mathbf{B}}$ must have the same size as expected from the additional information provided by the measurement:

$$\hat{\mathbf{B}} = (\mathbf{I}-\mathbf{GH})\mathbf{B}(\mathbf{I}-\mathbf{GH})^T + \mathbf{GRG}^T = (\mathbf{I}-\mathbf{GH})\mathbf{B}. \qquad (2.23)$$

The updated ensemble members directly correspond to an update of the state covariance $\hat{\mathbf{B}}_t$ and circumvent spurious underestimation of $\hat{\mathbf{B}}_t$.

To complement this mathematical explanation, and to provide more intuitive understanding, Appendix A.4 contains a toy implementation of an Ensemble Square Root Filter for estimating function parameters. It optimizes a simple function consisting of Gaussians with shifted centre to fit simulated measurements. The example in Fig. A.23 shows the limitation that the parameters need to be preconditioned to give changes in these parameters similar weight. CarbonTracker realizes this for real-world inversion by optimizing scaling factors applied to the prior flux. This follows the assumption that the relative average magnitude of the prior flux fits reality reasonably well, so that realistic adjustments to the fluxes should be proportional to the total magnitude of the prior. The assumption can cause problems in regions which get assigned a prior flux very close to zero, because this prior likely consists of opposing non-zero fluxes which add up to zero.



(a) (b)

Figure 2.12: Map of the ecoregions and covariance for their dimensionless scaling factors from CarbonTracker. Figure 2.12a is parsed from the regions.nc input file while Figure 2.12b is taken from the validation output of a CarbonTracker run with ecoregions. The rectangle areas before index 200 correspond to transcom regions in the southern hemisphere in which the ecoregions are coupled to compensate the lower density of measurements in these regions while the rectangle at the bottom right corresponds to the covariance of the ocean regions, estimated following Jacobson et al. (2007).

Since the flux adaptation is scaled by the magnitude of the prior flux, those regions cannot be adapted. The impact of this effect gets smaller with increasing resolution of the model and prior.

Overall, CarbonTracker's EnSRF approach requires running the transport model $\mathcal{H}$ for E ensemble members over the time period covered by all time steps t. At each time step t the transport model is sampled at all measurement instances within the time step and the above methodology is followed.

The choice of assimilation time window, here five weeks, implies that CarbonTracker can adjust surface fluxes only when their effects are observed at a site within five weeks of atmospheric transport. In the zonal direction this limitation is of little consequence, because typical global zonal transport time scales are on the order of weeks. But in meridional direction and especially for inter-hemispheric transport, where the transport time scales are on the order of months, this choice needs to be taken into account when interpreting flux results. The time stepping also defines the temporal binning of one-week fluxes.

The spatial binning of CarbonTracker's state vector follows the transcom regions (Gurney et al., 2000), further categorized into land regions with similar ecosphere following Olson et al. (1992) and ocean regions following the Ocean Inversion Fluxes (Jacobson et al., 2007) as described in the documentation of CarbonTracker North America.[10] In total, there are 240 flux ecoregions to be optimized, which is significantly less than the number of grid cells of the transport model operating on $6° \times 4°$ (longitude $\times$ latitude). The fluxes are further separated into 3 categories: biosphere/ocean, fire and fossil fuel. Only the category biosphere/ocean is optimized, the others

[10] CarbonTracker 2011_oi results and documentation are provided by NOAA ESRL, Boulder, Colorado, USA from the website esrl.noaa.gov/gmd/ccgg/carbontracker/CT2011_oi/. The site builds on the work from Peters et al. (2007).

are imposed from their priors following the assumption that fossil fuel fluxes are known with much higher precision than biosphere and ocean fluxes and that fire fluxes cannot easily be distinguished from biosphere fluxes, so they could not be interpreted separately. Altogether, temporal and spatial binning results in a state vector $\vec{x}_t$ with $240 \times 5 = 1200$ elements.

Optimizing scaling factors allows using a high resolution prior which stays in high resolution during the optimization, even though the state to be optimized has much fewer elements. Fig. 2.3 shows the optimized fluxes of a CarbonTracker run at $1^\circ \times 1^\circ$ resolution. The neighbouring source and sink pixels in tropical Africa clearly show the borders between ecoregions optimized by different scaling factors.

The structure of the background covariance $\mathbf{B}_t$ in the Northern Hemisphere is a diagonal matrix with a variance of 0.64 (80% standard deviation) in units of dimensionless flux scaling factors. In tropical and many Southern Hemisphere regions, the ecosystems are coupled with exponentially decreasing covariance, selected such that the total covariance in the transcom region matches the variance in Northern Hemisphere regions. The covariance for ocean regions uses the results of the ocean inversion by Jacobson et al. (2007). Fig. 2.12 shows the ecoregions and the spatial covariance of their flux scaling factors. Temporal covariance in CarbonTracker stems from processing observations multiple times in the time stepping approach. The observation covariance $\mathbf{R}$ is assumed diagonal.

The version of CarbonTracker used in chapter 3 is derived from version 1.0 of the code maintained by Wageningen University with the same state vector as CarbonTracker North America (as used in Peters et al., 2007) and without a zoom region.

The assimilation of total column measurements in chapter 4 uses the new "gridded" version of CarbonTracker Europe, developed by Wouter Peters. The system works like the setup by Babenhauserheide et al. (2015), except for the change that the ecoregions in the state vector are split on the resolution of a 50×50 km grid with covariance of the unitless scaling factors around 0.6 between the grid cells within each ecoregion. This recreates the ecoregion structure in the gridded state vector while providing more freedom for adjustments of the flux within a given ecoregion.

This "gridded" version is yet to be published, however, and lacks the amount of polish which went into the non-gridded version (used in Babenhauserheide et al., 2015; Peters et al., 2007).

2.5.4 Variational Method: TM5-4DVar

Whereas the EnSRF in CarbonTracker reduces the dimension of the minimization problem of Eq. (2.6) by solving sequentially for time-sliced state vectors, the 4DVar method in TM5-4DVar leaves the dimension of the state vector intact and approximates the solution using a limited set of search directions, corresponding to the dominant singular vectors (Golub and Reinsch, 1970) of the inverse problem to approach the minimum of the cost function step-by-step (Meirink et al., 2008; Basu et al., 2013). The iterative minimization of Eq. (2.6) in TM5-4DVar is described in detail by Chevallier et al. (2005) and Meirink et al. (2008). It employs the conjugate gradient algorithm (Navon and Legler, 1987) which is equivalent to the Lanczos method (Lanczos, 1950; Fisher and Courtier, 1995) and requires calculation of the cost function gradient

$$\nabla_{\vec{x}} J = \mathbf{B}^{-1}(\vec{x}_n - \vec{x}_\mathrm{b}) - \mathbf{H}^T\mathbf{R}^{-1}(\vec{y} - \mathbf{H}\vec{x}_n) \tag{2.24}$$

where subscript n indicates the nth iterative step. The adjoint formulation of TM5 allows calculating the cost function gradient by a single run of the transport model and its adjoint

(Errico, 1997; Chevallier et al., 2005). The conjugate gradient algorithm further provides the leading eigenvalues and eigenvectors of the preconditioned Hessian

$$\boldsymbol{\nabla}_{\vec{\chi}}(\boldsymbol{\nabla}_{\vec{\chi}}J) = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}, \tag{2.25}$$

which is the second derivative of the cost function J with respect to the dimensionless preconditioned state $\vec{\chi}$ defined as $\vec{x} = \mathbf{L}\vec{\chi} + \vec{x}_\mathrm{b}$, where $\mathbf{L}$ is the preconditioning matrix with $\mathbf{B} = \mathbf{L}\mathbf{L}^T$. This can be used to construct the inverse of the state covariance $\hat{\mathbf{B}}^{-1}$ as defined in Eq. (2.9). After n steps, corresponding to n runs of the forward and the adjoint model, the minimization algorithm yields an optimized state estimate $\hat{\vec{\chi}}_n$ and the first n eigenvalues λ_i ($\lambda_i > 1$) and eigenvectors $\vec{v}_i$ ($i = 1, \ldots, n$) for the eigensystem of the preconditioned Hessian. The latter can be used to construct an approximate error covariance matrix,

$$\hat{\mathbf{B}}_n \approx \mathbf{B} + \sum_{i=1}^{n}\left(\frac{1}{\lambda_i} - 1\right)(\mathbf{L}\vec{v}_i)(\mathbf{L}\vec{v}_i)^T. \tag{2.26}$$

With increasing number of iterations, the optimized state vector $\hat{\vec{x}}_n$ approaches the optimal state vector $\hat{\vec{x}}$ at the minimum of the cost function and the approximate state covariance $\hat{\mathbf{B}}_n$ approaches $\hat{\mathbf{B}}$ from above, so that the estimated uncertainty is always larger than the analytical value (Basu et al., 2013). For practical purposes the iteration is stopped when the gradient norm reduction exceeds a threshold, i.e.

$$|\boldsymbol{\nabla}_{\vec{x}}J(\vec{x}_n)| \leq \eta \cdot |\boldsymbol{\nabla}_{\vec{x}}J(\vec{x}_0)| \tag{2.27}$$

with the constant chosen to be $\eta = 10^{-9}$ here. As shown in Fig. 2.13, the descent is roughly exponential. An increase from $\eta = 10^{-6}$ to $\eta = 10^{-9}$ roughly increases the number of iterations from 60 to 90 in our study.

TM5-4DVar's state vector $\vec{x}$ is binned temporally in monthly fluxes and spatially on the transport model grid scale, i.e. 6° ×4° longitude × latitude. Fig. 2.14 shows yearly optimized flux for the baseline run. Since these grid boxes are coarser than the mapping of ecoregions in CarbonTracker, the results from both systems can only be compared on transcom regions which form the smallest exact superset of both structures of the state vectors. Fluxes are categorized into biosphere, ocean, fire and fossil fuel. To create a setup comparable to CarbonTracker, only biosphere and ocean fluxes are optimized. The background covariance $\mathbf{B}$ of the state vector is characterized by a global temporal and spatial correlation length. By default TM5-4DVar uses an exponential decay with a temporal and spatial length scale of 1 month and 200 km for biosphere fluxes and 3 months and 1000 km for ocean fluxes. As such, the temporal binning of TM5-4DVar's state vector containing monthly bins is about a factor 4 coarser than the temporal binning of CarbonTracker's weekly bins. TM5-4DVar's spatial binning has a different overall structure. Whereas CarbonTracker's prior fluxes are fully correlated inside the 240 ecoregions and mostly uncorrelated between different ecoregions, the correlation of TM5-4DVar's fluxes falls off exponentially around each grid box. The exponential decay in TM5-4DVar's temporal background correlation limits the effects of observations in time. However, TM5-4DVar has no strict limit on the time window during which observations can be linked to fluxes but rather reduces the strength of the influence with temporal lag. TM5-4DVar can adjust surface fluxes in response to any observation during the entire considered time period given that the transport model reveals a link between fluxes and observations. As for CarbonTracker, the observation covariance $\mathbf{R}$ is assumed diagonal.

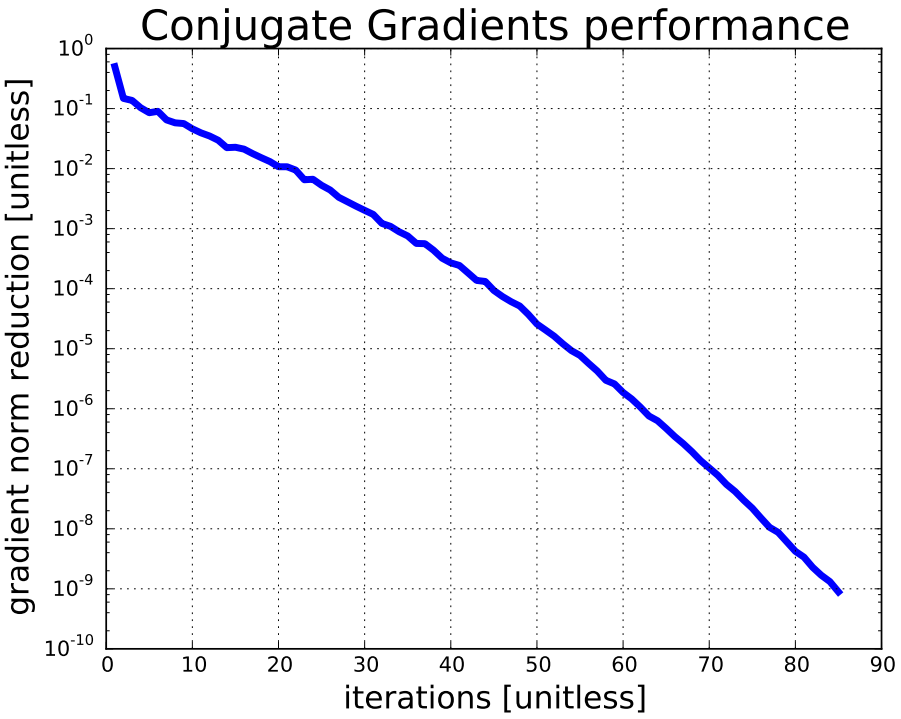


Figure 2.13: Gradient norm reduction of TM5-4DVar by iteration. 90 iterations take about one week on our 6-core desktop system. This is on the same order of magnitude as CarbonTracker runs using 8 CPU cores.

2.6 The promise of assimilating total column measurements

The atmospheric concentration of CO_2 is controlled by sources and sinks on the surface (anthropogenic and biospheric fluxes) and atmospheric transport. As such, good knowledge of the atmospheric concentration and the transport allows reconstructing sources and sinks. However, in-situ measurements suffer from a strong influence of local fluxes.

In contrast to the point-like in-situ measurements, column data provides information about the entire atmospheric column. Due to the limited vertical mixing speed with transport times of about a week to a month from the ground to the tropopause (Jacob, 1999), column measurements contain information about the fluxes in distant regions and complement in-situ measurements which provide highly accurate information about the CO_2 concentration at a given height. A column measurement and an in-situ measurement at the same position contain information which part of a change in concentration originates from their immediate vicinity and which part originates from farther away. Column measurements from the ground give high-frequency measurements at a fixed point, while satellite data gives good coverage of measurements throughout the world with limited temporal resolution.

There are several studies which try to use total column measurements from satellite missions to provide better global coverage and extend the measurements above the boundary layer (i.e. Basu et al., 2013), but these suffer from changing observation coverage due to cloud patterns and

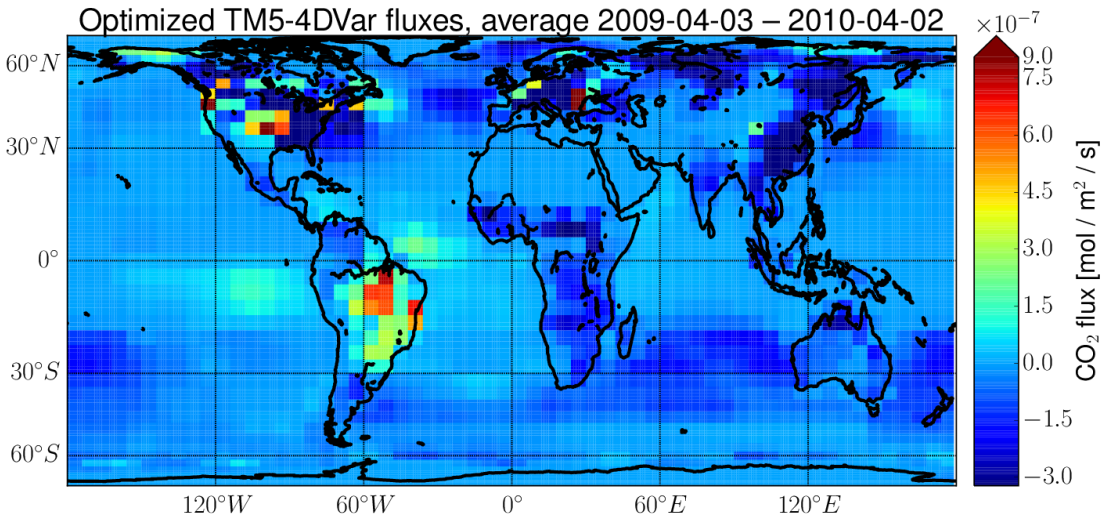


Figure 2.14: 6x4° longitude×latitude map of biosphere+ocean fluxes for TM5-4DVar from April 2009 to April 2010. Shown for illustrative purposes. Maps of CarbonTracker and TM5-4DVar cannot be compared on their native resolution due to incompatible structure of the underlying state vector.

location dependent systematic biases (Chevallier, 2015) which are hard to correct with the current sparse availability of validation aircraft missions and the short average lifetime of satellites.

To avoid the complications from satellite measurements, this study extends and complements the observation network with total column measurements from the TCCON network. It follows prior work from Chevallier et al. (2011) who investigated the viability of assimilating total columns as alternative to in-situ measurements. Here, however, the total column measurements are used together with in-situ measurements and implemented in the well established CarbonTracker system. Properties of column measurements are described in section 2.2.2, while section 4.1 shows the approach taken in this study to assimilate total column measurements.

In chapter 4 this study investigates surface fluxes estimated from 4 years of inversion (2009–2012), using columns from TCCON and in-situ measurements from obspack (NOAA Environmental Sciences Division, Oak Ridge National Laboratory, 2013) together in the established CarbonTracker data assimilation framework. This extends the work from Chevallier et al. (2011) by running a shared assimilation of TCCON and in-situ, by estimating the representativeness error of TCCON measurements and by investigating how assimilating total column measurements affects model-estimated European fluxes.

Chapter 3

Uncertainty of CO_2 surface fluxes constrained by in-situ observations

Comparing two different inverse models used for real-world flux assimilation allows estimating the systematic errors due to the choice of the approximations in the models. Varying the assimilated observations shows the dependence of systematic errors on the density of observations and points to a way for minimizing the flux uncertainty.

This part of the study was conducted in collaboration with Wouter Peters, one of the authors of CarbonTracker, and Sourish Basu and Sander Houweling, two of the authors of TM5-4DVar. It was published in Atmospheric Chemistry and Physics (Babenhauserheide et al., 2015).

3.1 Comparing CarbonTracker and TM5-4DVar

We estimate the uncertainty of fluxes from two models used for real-world flux estimation by running them with harmonized input values. We compare their modelled concentration fields with non-assimilated measurements to get a measure of the quality of their results. Then we compare their optimized fluxes against each other to identify how their flux attribution differs. When they provide concentration fields of indistinguishable quality, the differences between their optimized fluxes gives an estimate of the uncertainty in the fluxes which stems from differences in their implementation (i.e. inverse method and flux parametrization).

The two methods are CarbonTracker and TM5-4DVar, as described in section 2.

Both methods are used in a number of studies. CarbonTracker studies include estimates of global CO_2 fluxes (Peters et al., 2007, 2010), European fluxes (Meesters et al., 2012), Asian fluxes (Zhang et al., 2014) as well as ^{13}C isotope studies (van der Velde et al., 2014). Studies with TM5-4DVar include CO_2 flux estimation (Basu et al., 2013), CO estimation (Hooghiemstra et al., 2011) and CH_4 emission estimates (Meirink et al., 2008; Bergamaschi et al., 2010; Houweling et al., 2014). Additionally the models were employed in several multi-model comparison studies (e.g. Schulze et al., 2009; Peylin et al., 2013; Thompson et al., 2014).

To analyse the data assimilation (DA) systems, we focus on a case study for the period from 2009 to 2010 and use observational constraints collected by an in-situ measurement network and compiled by the NOAA Environmental Sciences Division, Oak Ridge National Laboratory (2013, exact version: obspack PROTOTYPE v1.0.2 2013-01-28). CarbonTracker and TM5-4DVar are examples of an ensemble method and a variational method: EnSRF and 4DVar. Besides the

mathematical treatment of the inversion, CarbonTracker and TM5-4DVar differ in the design of the state vector. CarbonTracker and TM5-4DVar are introduced in sections 2.5.3 and 2.5.4.

Our goal is to evaluate the impact of the inverse method (including the flux representation) on the estimated surface fluxes. Therefore, we must ensure that the other components of the DA systems – the observations to be assimilated, the transport model and the prior assumptions – are the same.

3.1.1 Harmonizing the inverse methods

DA approaches differ in four main characteristics: first, they ingest different observational constraints, for example in-situ concentration measurements at different sites. Second, they represent sources and sinks of carbon differently, for example by binning them by vegetation type or on a latitude/longitude grid. Third, they relate sources and sinks to observed atmospheric abundances using different air-mass transport models (their impact on fluxes was estimated by Gurney et al., 2004). And fourth, they use different inverse methods that find the best estimate of the source-sink distribution using the transport model, the observational constraints, the representation of sources and sinks and a prior estimate of the sources and sinks. Differences in these characteristics contribute to the differences in flux estimates from different studies. To analyse the impact from the representation of sources and sinks and from the inverse method, it is therefore necessary to harmonize the observational constraints, the transport model and the prior concentration, flux and flux covariance estimates between the approaches which are compared.

That allows us to actually separate model performance from the quality of the input data. As input data the models use the observations $\vec{y}$ (real measurements, see section 2.2), a transport model $\vec{H}$ with meteorological data $\vec{M}$ (the TM5 model using the era interim data-set from the ECMWF weather model, section 2.4), prior fluxes $\vec{x}$ (from several biosphere models and an ocean model, section 2.3), an initial CO_2 concentration field $\vec{C}$ (from a prior run of CarbonTracker).

In our approach, we adjust the input for TM5-4DVar to fit the CarbonTracker inputs. In the following sections we describe the input parameters and data of the methods and how we harmonize them.

3.1.2 Transport model and observation operator

To connect concentration measurements and surface fluxes, CarbonTracker and TM5-4DVar use a transport model which transports the CO_2 tracer using meteorological fields. Both models use the Tracer Model 5 (TM5) as described by Krol et al. (2005) which utilizes meteorological data from the European Centre for Medium-Range Weather Forecasts (ECMWF, 2013). To reduce discretization errors, TM5 uses the gradient between the concentrations in adjacent grid boxes (the slopes scheme Russell and Lerner, 1981), This scheme is also used to interpolate concentrations for comparison to measurements with sensitivity on scales below the transport model scale.

For CarbonTracker, we follow the setup used by Peters et al. (2007). For TM5-4DVar our setup differs from the setup used by Basu et al. (2013) in one main aspect to be consistent with CarbonTracker: the CO_2 concentration field for in-situ ground sites is sampled in the second vertical model layer ($\approx$ 980hPa $\approx$ 170m) or higher instead of in the first model layer ($\approx$ 994hPa $\approx$ 50m) or higher. Except for these adjustments and some minor differences due to different interfaces of the inverse methods, the versions of TM5 used by the CarbonTracker and TM5-4DVar systems we are using are the same.

3.1.3 Background flux and initial guess

CarbonTracker and TM5-4DVar use the same background fluxes and initial concentration fields. The background fluxes used in both models are described in section sec{sec:background-flux}.

The initial concentration field is generated from the output of a previous CarbonTracker run which ended on 1 January 2007. The field for 2009 is derived by increasing the concentration by 1.9 parts per million (ppm) per year. The value 1.9 ppm was chosen based on tests of the fit to observation sites in the first month of 2009.

The covariance of the fluxes is defined in the models as described in sections 2.5.3 and 2.5.4. We harmonize the overall covariance by adjusting the prior flux uncertainty in TM5-4DVar to 172.59% of the flux for ocean grid boxes and to 199.17% for land grid boxes to match uncertainty of a CarbonTracker run with a monthly cycle for global and continental aggregates.

3.1.3.1 Background flux covariance harmonization process

Since CarbonTracker does not have a temporal flux uncertainty aggregation scheme which takes correlations into account, its background (prior) flux uncertainty cannot simply be scaled to fit the uncertainty of TM5-4DVar (or vice versa). The only uncertainties which stem directly from the inverse method are on weekly scale, more exactly: on the scale of a cycle length. These uncertainties are calculated from the spread of the ensemble runs and as such they are guaranteed to correspond to the flux covariance. Getting a quantity which can be compared is possible with a harmonization run of CarbonTracker in which its cycle length is increased to 30 days – corresponding to the one-month temporal flux binning of TM5-4DVar. With this 30 days covariance data for CarbonTracker we then adjusted the flux uncertainty parameters of TM5-4DVar to minimize the mismatch between the flux uncertainty of CarbonTracker and TM5-4DVar on global scale and on the scale of transcom regions. The chosen values 172.59% of the flux for ocean grid boxes and 199.17% for land grid boxes strike a balance between good global and good regional match.

Due to the different ways of specifying the state vector $\vec{x}$ and its covariance $\mathbf{B}$ in CarbonTracker (weekly with ecoregions) and TM5-4DVar (monthly gridded with global covariance parameters), it is not possible to get an exact match of the flux uncertainties. This is a result of comparing real-world systems used for flux estimation to not only capture theoretical effects but also differences which show in practical use. While making the comparison more complex, this choice allows getting a better understanding of the uncertainties due to the large amount of implementation decisions which have to be taken for a production system. The remaining mismatches in the prior flux uncertainty can have an effect on the estimated fluxes. This effect has to be taken into account for interpreting a posteriori flux differences. Section 3.2.2.1 includes an example of such an analysis. The remaining mismatches in the flux uncertainty per transcom region and month are provided in Appendix A.1.

3.1.4 Observations and observation errors

Both DA systems use the same observations from the "obspack" compilation of in-situ CO_2 concentration measurements (Masarie et al., 2014; NOAA Environmental Sciences Division, Oak Ridge National Laboratory, 2013, version: PROTOTYPE v1.0.2 2013-01-28), as described in section 2.2.1. Discrete (e.g. one sample per week) measurements from surface flask sites, in situ continuous (and semi-continuous) measurements from surface sites and towers, and aircraft campaign measurements are collected, aggregated and quality screened to make them suitable for

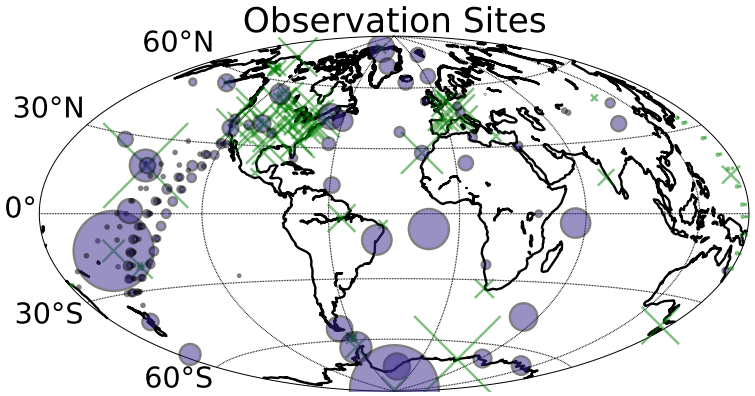


Figure 3.1: Ground based in-situ observation sites assimilated by the baseline run (circles) and non-assimilated validation sites (crosses). The radius of a circle or cross is proportional to $\frac{\sqrt{N}}{\sigma}$ with N the number of available measurements at the site in this location and σ their uncertainty estimate. The larger the symbol, the larger the weight of the site in the model.

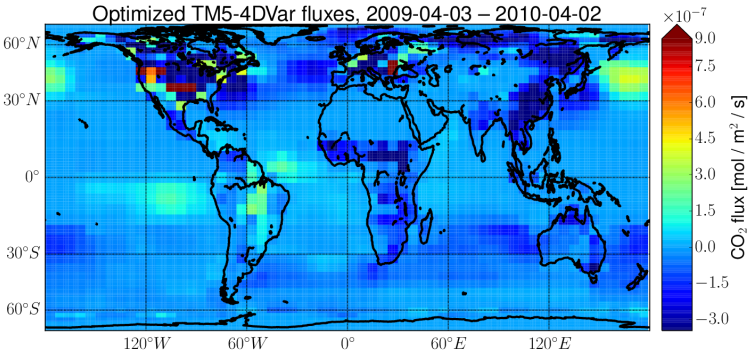


Figure 3.2: 6°×4° longitude×latitude map of fluxes for TM5-4DVar aggregated from April 2009 to April 2010 when assimilating all observations. Due to the strong gradients between grid boxes in North America and the source in the northern Pacific we exclude in the baseline run 5 sites from the assimilation which have more than 1000 observations in the time period studied.

inverse flux estimation. For our baseline CarbonTracker and TM5-4DVar runs, we exclude 21 measurement sites from the assimilation to use them as validation sites.

Additionally we take out 5 sites which have more than 1000 measurements in the assimilation period. This is to keep the TM5-4DVar results representative of TM5-4DVar runs which use the native TM5-4DVar input. When using these 5 sites with the CarbonTracker preprocessing, TM5-4DVar shows strong gradients between neighbouring grid cells in North America and a source in the northern Pacific (see Fig 3.2) which it does not show when processing its native set of observations. In addition to these 26 excluded sites, there are 24 further sites from which the default run of CarbonTracker uses no data or only a subset of the observations. Reasons for not using some of the observation data of a site include that the data is assumed not representative of its grid-cell or recorded in aircraft campaigns. Figure 3.1 shows the sites assimilated or excluded.

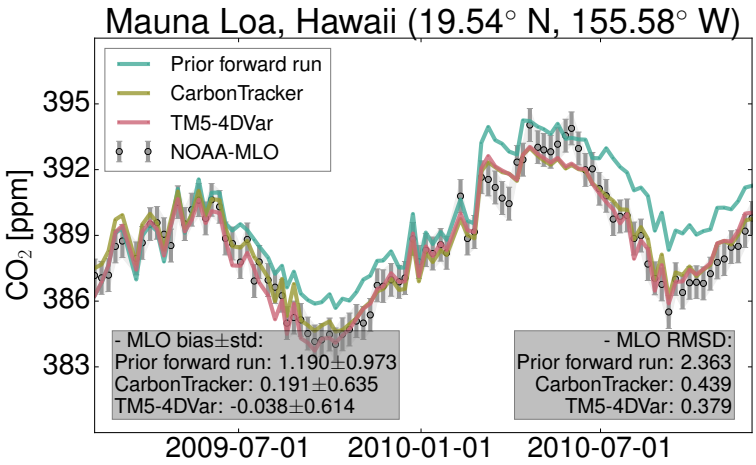


Figure 3.3: Time series of measured and modelled CO_2 concentrations from CarbonTracker and TM5-4DVar at Mauna Loa, Hawaii, Pacific (assimilated weekly flasks), NOAA site code MLO. Also shown are the concentrations for obtained from a forward run of the transport model using the a priori background flux estimates.

In our setup CarbonTracker and TM5-4DVar use the same representativeness errors, as described in section 2.2.1.

3.2 Uncertainties due to model design

3.2.1 A posteriori concentration fields

As a first step, we compare and validate the performance of CarbonTracker and TM5-4DVar by evaluating the difference between measured and modelled CO_2 concentration fields at the location of various ground sampling stations. Comparing concentration fields at the assimilated sites provides a check to verify that data assimilation works in both systems. Comparing measured and modelled concentrations at non-assimilated sites demonstrates to what extent the data assimilation approaches yield improvements where observational constraints are distant in space and/or time. CarbonTracker and TM5-4DVar are both run with the baseline setup (as described in Sect. 3.1.1) for a 23 month period starting on 1 February 2009.

3.2.1.1 Assimilated sites

As an example for an assimilated site, Fig. 3.3 shows a time series of measured and modelled CO_2 concentrations at Mauna Loa (MLO), Hawaii, located 3399 meter above sea level (masl) in the Pacific. For the period from 1 February 2009, to 30 December 2010, the models assimilate 94 weekly flask measurements. We compare the observations to a posteriori and a priori model concentrations. The a posteriori concentrations are sampled using the a posteriori surface fluxes estimated by CarbonTracker or TM5-4DVar. The prior model concentrations are sampled using the background (prior) flux estimate common to both models.[1] The Mauna Loa record

[1]For CarbonTracker this requires a run which does not assimilate any observations. Or, in practical terms, with the observation uncertainty set to values on the order of 1000 ppm. The mismatch between observations and model are mostly below 10 ppm and the uncertainty of fluxes is at roughly 80% of the flux, so this ensures

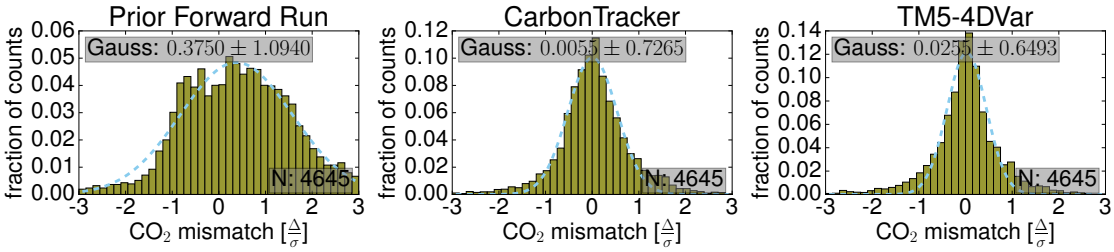


Figure 3.4: Histograms of the mismatch between measured and modelled CO_2 concentrations for all assimilated measurements using prior fluxes, CarbonTracker optimized fluxes and TM5-4DVar optimized fluxes. The histograms show residuals for one year (3 April 2009 to 2 April 2010) which are normalized by the estimated representativeness error. The line on top of the histograms is a fit of a Gauss function to the histogram. The parameters in the top left show the bias and standard deviation of the Gaussian. The bottom right shows the number of measurements which were accumulated into the histogram.

demonstrates that the a posteriori concentrations produced by both models match the observations within the uncertainty estimate and that the match is substantially better than for the prior concentration fields. Differences between CarbonTracker and TM5-4DVar are much smaller than the representativeness error of the measurements at Mauna Loa (0.75 ppm) over the entire period. This is consistent with the results at other sites.

The mismatch between measured and modelled CO_2 concentrations for all assimilated measurements is shown in Fig. 3.4, with the prior concentrations, the a posteriori concentrations optimized by CarbonTracker, and the a posteriori concentrations optimized by TM5-4DVar. The concentration mismatch is normalized by the representativeness error of the observations such that a (unitless) mismatch of 1 corresponds to a mismatch with the magnitude of the representativeness error. Unlike the time series for Mauna Loa, the histograms only integrate over the 1 year period 3 April 2009 to 2 April 2010 in order to be consistent with the analysis of the a posteriori surface fluxes in Sect. 3.2.2. This time period gives the models sufficient spin-up and spin-down time, because the initial concentration is already well-optimized by a previous CarbonTracker run.[2]

The concentrations from the Prior Forward Run in Fig. 3.4 reveal an overall bias in the normalized (unitless) mismatch of 0.37 with a standard deviation of 1.09. Tentatively, the prior fields show a dipole pattern with peaks around -1 and 1 which can be traced back to the Northern Hemisphere prior generally overestimating the observations and the Southern Hemisphere prior generally underestimating the observations. The CarbonTracker and TM5-4DVar histograms show small biases of 0.006 and 0.025 with a standard deviation of 0.727 and 0.650, respectively. Compared to the prior, both DA systems improve the overall bias and they substantially reduce the spread of the observation-model mismatch. Normalized standard-deviations smaller than 1 indicate that the mismatch is on average smaller than the estimated representativeness error, which points to a conservative choice of representativeness errors and consequently a stronger than optimal influence of the prior flux estimate. However, avoiding this would require using the output of the assimilation systems to adjust their input parameters which could lead to transient errors in the result.

that the observations have negligible impact on the optimized fluxes while still being sampled as usual. The prior output of Carbontracker is only the prior of the current time step, but builds on the optimized concentration field from the previous step.

[2] If there were significant differences in the concentration, this assumption would have to be re-evaluated.

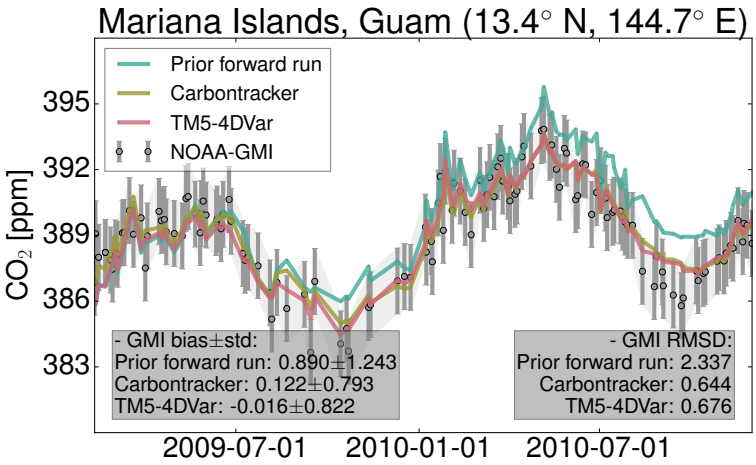


Figure 3.5: Time series of measured and modelled CO_2 concentrations from CarbonTracker and TM5-4DVar at Guam, Mariana Islands, Pacific (non-assimilated). Also shown are the concentrations obtained from a forward run of the transport model using the a priori background flux estimates (Prior forward run).

The histograms for a posteriori concentrations of CarbonTracker and TM5-4DVar reveal some non-Gaussian behaviour with long tails toward greater mismatch and with a narrow peak at the centre. The tails most likely stem from temporally varying contributions to the representativeness error which our input data assumes constant in time. The narrow peak likely stems from two sources: first, sites with high frequency measurements are assumed to provide uncorrelated data in the models and as such give a stronger constraint than sites with low frequency measurements. Second, an already well-optimized prior which is close to the observations causes the models to stick to the prior in a sparse observation network.

In summary, both models show similar performance for assimilated sites, and the assimilation substantially reduces the mismatch between modelled and measured concentrations at assimilated sites.

3.2.1.2 Non-assimilated sites

Next, we evaluate the performance of the DA systems for sites whose observations are not assimilated. These sites provide independent validation of the results. Figure 3.5 shows a time series of flask measurements in Guam on Mariana Islands (GMI), West Pacific. In contrast to Mauna Loa, the measurements are taken at sea level, and are not assimilated by the CarbonTracker and TM5-4DVar inverse models. The observation error in Guam is 1.5 ppm, and the modelled concentrations agree well with measurements taken at the site. Both, CarbonTracker and TM5-4DVar, reproduce the measurements similarly well with a respective bias of 0.12 and 0.02 ppm. Their standard deviation of 0.79 and 0.82 ppm, is greater than the standard deviation at Mauna Loa, our selected example for assimilated sites. The prior concentrations on the other hand deviate substantially from the measurements, with a bias and standard deviation of 0.89 and 1.24 ppm, respectively.

The histograms of model-observation mismatch are shown in Fig. 3.6, for the concentrations of a Prior Forward Run and for the a posteriori CarbonTracker and TM5-4DVar runs. They show a

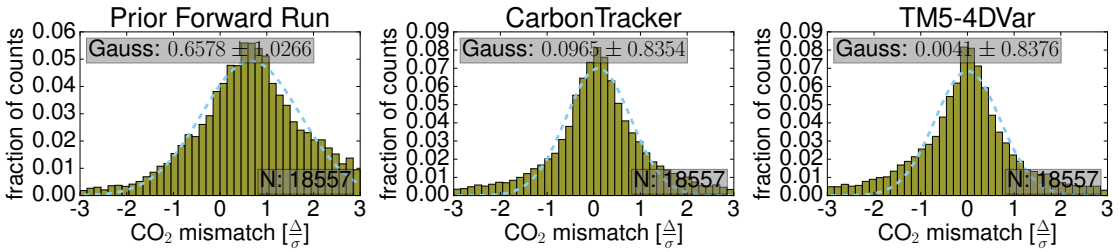


Figure 3.6: Histograms of the mismatch between measured and modelled CO_2 concentrations for all non-assimilated samples using prior fluxes, CarbonTracker optimized fluxes and TM5-4DVar optimized fluxes. The histograms show residuals for one year (3 April 2009 to 2 April 2010) which are normalized by the estimated representativeness error. The line on top of the histograms is a fit of a Gauss function to the histogram. The parameters in the top left show the bias and standard deviation of the histogram. The bottom right shows the number of measurements which were accumulated into the histogram.

higher number of measurements than the histograms of assimilated sites, because many of the non-assimilated measurements come from continuous sampling sites and aircraft campaigns which provide a high number of measurements. Normalized bias and standard deviation of the prior mismatch aggregated for all sites are 0.66 and 1.03, respectively. The normalized biases of the mismatch for CarbonTracker and TM5-4DVar are 0.097 and 0.004, respectively, and the standard deviation of the histograms are 0.835 and 0.839, indicating that assimilating observations with the DA systems substantially improves the match to independent data when compared to the prior performance. The spread of the a posteriori model-observation mismatch, however, is somewhat greater than for the comparison to assimilated measurements. This is as expected and indicates a slightly worse performance of both methods for the non-assimilated than for assimilated sites.

3.2.1.3 Robustness of the result

CarbonTracker a posteriori concentrations show a larger bias for non-assimiliated measurements (0.097) than for assimilated measurements (0.006). TM5-4DVar biases are more similar for non-assimiliated (0.004) and assimilated measurements (0.025). In order to investigate whether these differences are likely to be an artefact of our selection of validation sites, we conduct a resampling experiment. Out of the 50 sites for which there are non-assimilated observations – our 26 validation sites, aircraft measurements and sites for which only a given measurement method is assimilated – we randomly select subsets of 25 sites and recalculate the statistical model-observation bias for non-assimilated measurements. Then we repeat the exercise 9 times and examine the distribution of the resampled CarbonTracker and TM5-4DVar biases. Figure 3.7 shows that the normalized biases for the CarbonTracker baseline run consistently scatter around 0.08 with a standard deviation of 0.04 while the TM5-4DVar average bias and standard deviation are -0.04 and 0.07, respectively.

So, while CarbonTracker a posteriori concentrations appear offset from the (non-assimilated) observations, TM5-4DVar does not show a significant overall bias but greater station-to-station variability for the model-observation mismatch.

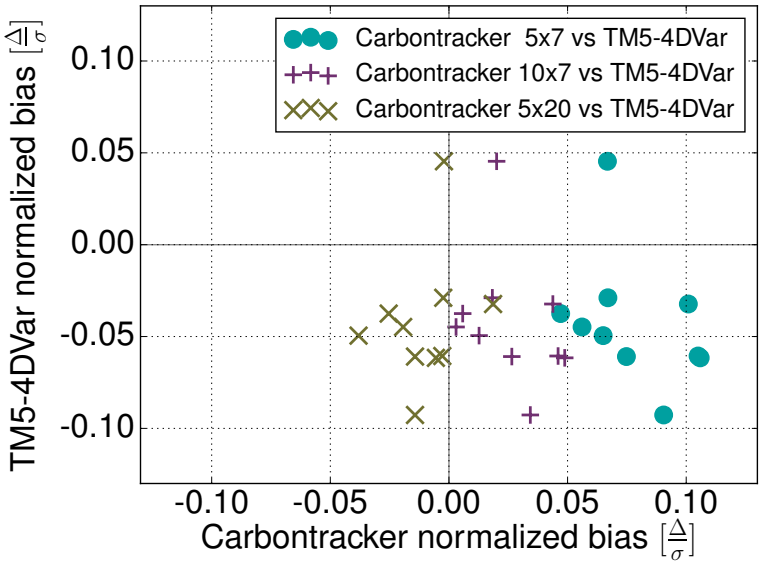


Figure 3.7: Model-measurement bias of TM5-4DVar against CarbonTracker for non-assimilated measurement sites. Each symbol corresponds to a case resampling exercise where the biases are calculated for 25 randomly drawn sites out of the total 50 resampling sites listed in Table A.1. The baseline run (dots) is compared to a CarbonTracker run with the assimilation period extended to 5×20 days ($\times$) instead of 5×7 days and 10×7 days ($+$).

3.2.1.4 Impact of the CarbonTracker assimilation window length

In order to investigate whether the robust bias our resampling found for CarbonTracker can be due to the choice of the EnSRF assimilation time window, we vary CarbonTracker's lag and cycle parameters. Figure 3.8 illustrates the effect of the window length on the model-observation mismatch at Syowa (SYO), Antarctica. Syowa is located far from any major sources or sinks to be adjusted by the DA systems. Therefore, the DA systems cannot match the Syowa measurements by flux adjustment unless they account for long-reaching correlations between concentrations and fluxes. While TM5-4DVar allows for such connections, CarbonTracker's baseline assimilation window strictly limits these to 5 weeks, which is shorter than the transport time scales from strong flux regions to Antarctica. Therefore, the baseline CarbonTracker run shows a small but systematic underestimation of the CO_2 concentration by up to 0.5 ppm observed in Syowa in summer and fall 2009 while TM5-4DVar a posteriori concentrations match well (not shown). Increasing or decreasing CarbonTracker's assimilation window length respectively improves or deteriorates the match to Syowa observations, showing that the assumed temporal correlations play a role. For sites which are closer to biosphere regions, this effect could manifest as flux mi-sattribution, which would yield a mismatch to non-assimilated stations.

Figure 3.7 illustrates the resulting biases for our resampling assessment when CarbonTracker is run with an assimilation window of 10×7 days or 5×20 days instead of 5×7 days. For 10×7 the average normalized bias reduces to 0.03 with a standard deviation of 0.03 and for 5×20 the average normalized bias reduces to -0.01 with a standard deviation of 0.03. Both are consistent with TM5-4DVar's performance and better than the run with 5×7 days. This suggests that a longer assimilation window adds valuable information to CarbonTracker's DA system. It is unclear, though, whether this improved match to validation measurements translates

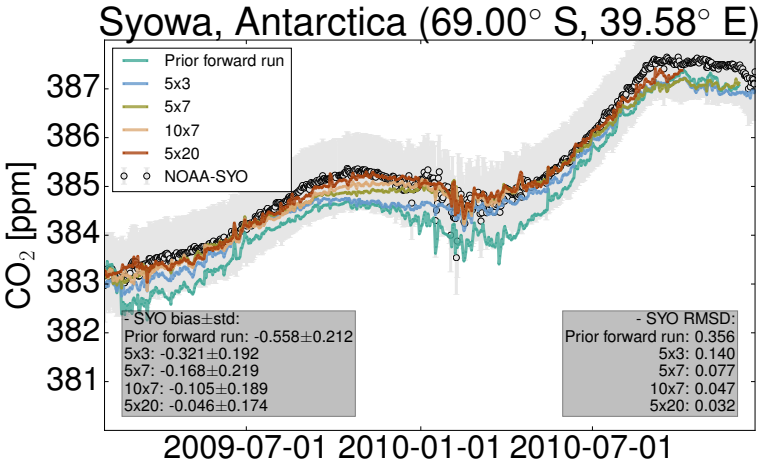


Figure 3.8: Time series of measured and modelled CO_2 concentrations in Syowa, Antarctica, for CarbonTracker runs with different length of the assimilation time window. The baseline run uses an assimilation window of 5×7 days. Colour coding of shorter and longer assimilation windows follows the legend (lag $\times$ cycle in days).

into improved flux estimates since transport model errors might have a larger impact for the longer assimilation windows. In section 3.2.2 we discuss additional effects from a larger bin size which may make a long assimilation window undesirable, despite the better match to validation measurements.

3.2.2 Comparison of a posteriori surface fluxes

Section 3.2.1 shows that the methods are of similar quality when comparing the a posteriori concentrations with assimilated and non-assimilated observations. Here, we turn to evaluating the a posteriori surface fluxes delivered by CarbonTracker and TM5-4DVar.

As first step we describe the results of the baseline runs. Then we analyse detectable features and the effect of a longer assimilation window in CarbonTracker.

Table 3.1: Yearly global CO_2 fluxes and uncertainty (standard deviation) from the Prior forward run and from the baseline runs of TM5-4DVar and CarbonTracker.

	Biosphere + Ocean	Uncertainty
Prior forward run	$-5.34\,\mathrm{Pg\,C\,a^{-1}}$	$1.86\,\mathrm{Pg\,C\,a^{-1}}$
TM5-4DVar	$-6.69\,\mathrm{Pg\,C\,a^{-1}}$	$1.07\,\mathrm{Pg\,C\,a^{-1}}$
CarbonTracker*	$-6.76\,\mathrm{Pg\,C\,a^{-1}}$	N/A

* CarbonTracker provides uncertainties on weekly scale. Aggregating them to yearly scale is not clearly defined and would not be comparable to TM5-4DVar.

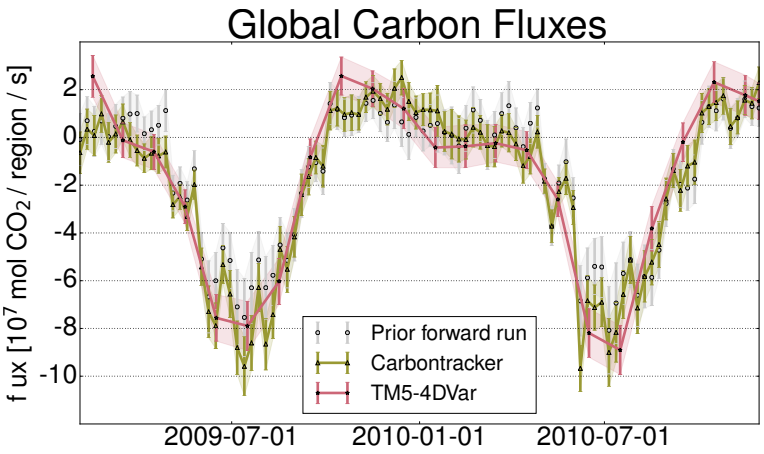


Figure 3.9: Global fluxes from the baseline runs of TM5-4DVar and CarbonTracker. The Prior is shown in the binning of CarbonTracker. The uncertainties shown for Carbontracker are aggregated spatially but not temporally. As such they represent the uncertainty of the estimated fluxes, calculated directly from the ensemble. These uncertainties are excluded from the annually aggregated graphs, because there is no method for temporally aggregating the uncertainties in a way which is comparable to the uncertainties estimated by TM5-4DVar.

3.2.2.1 Surface fluxes of the baseline run

For the baseline CarbonTracker and TM5-4DVar runs, Table 3.1 shows the globally aggregated a posteriori fluxes for the biosphere and oceans from 3 April 2009, to 2 April 2010. CarbonTracker and TM5-4DVar estimate a global carbon sink (due to the biosphere and oceans) which is stronger than the prior estimate by 1.42 and 1.35 PgC a^{-1}, respectively. We only show the uncertainty for the prior and TM5-4DVar which is calculated as described by Basu et al. (2013) and in section 3.1.3.1, because for CarbonTracker the aggregation of uncertainties from weekly to yearly scale requires using assumptions about the temporal correlation of the uncertainties. Due to these assumptions, the yearly uncertainties of TM5-4DVar and CarbonTracker would not be comparable, even if we adopted existing schemes as for example the one employed by Peters et al. (2005). The differences in the uncertainties would not be representative of actual differences in the models. Therefore we use the uncertainties from TM5-4DVar as a metric for comparisons. Different from the Monte-Carlo based uncertainty calculation which Chatterjee and Michalak (2013) used, the error propagation employed in TM5-4DVar always approaches uncertainties from above: the aggregated errors are larger than the analytical uncertainties at the exact minimum of the cost function (Personal Communication with Sander Houweling).

Due to this we expect our uncertainties to overestimate the real uncertainties from measurement and representativeness errors. With this caveat, the sink estimates of the two models are consistent within the TM5-4DVar uncertainties and also match previous findings for CarbonTracker (Peters et al., 2007). Examining the time series of globally aggregated surface fluxes in Fig. 3.9 confirms that the two DA systems are consistent on the global scale, both showing stronger summer uptake than the prior.

Figure 3.10 illustrates the a posteriori biogenic and oceanic fluxes aggregated over the one-year time period on continental scale regions. Agreement between CarbonTracker and TM5-4DVar is found for North America, Africa, Europe, and Australia, as well as for all the oceans except

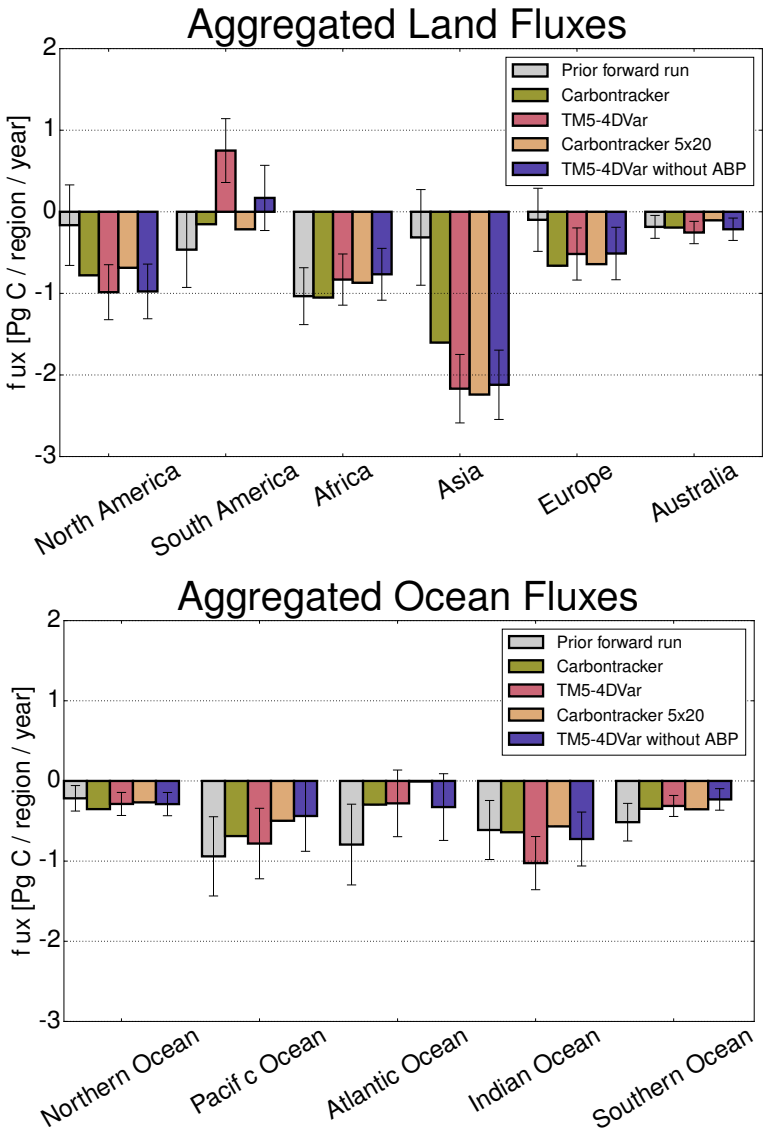


Figure 3.10: Fluxes from TM5-4DVar and Carbontracker aggregated on continental scale. The uncertainties for TM5-4DVar are calculated following Basu et al. (2013). The error bars for the prior are taken from TM5-4DVar. We show no uncertainties for CarbonTracker, because the aggregation of uncertainties from weekly to yearly scale is not clearly defined.

for the Indian Ocean. The optimized fluxes in these regions differ by less than the yearly uncertainties estimated from TM5-4DVar's statistical error aggregation (see Basu et al., 2013). On the other hand, the modelled fluxes from CarbonTracker and TM5-4DVar differ by more than their uncertainty in South America, Asia and the Indian Ocean. In South America they differ by roughly two times the estimated uncertainty, therefore we take a more detailed look at this discrepancy.

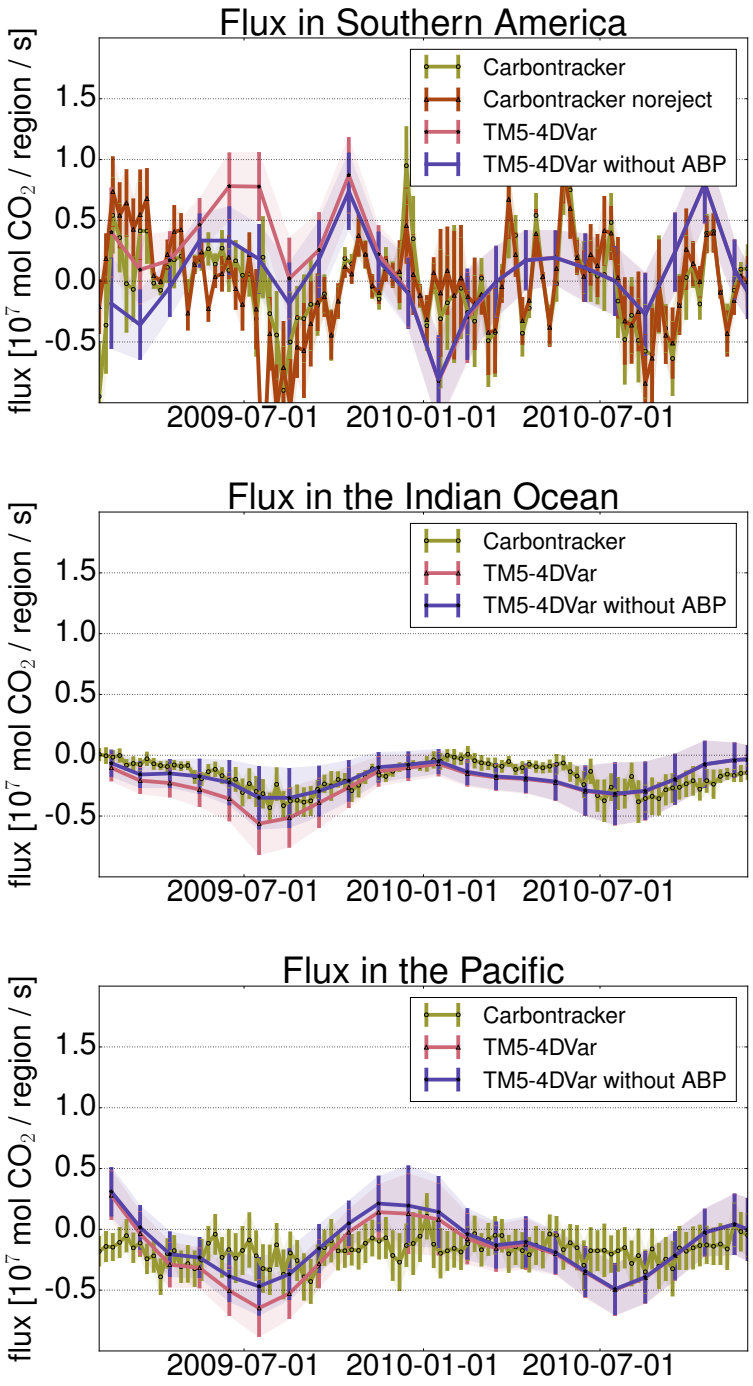


Figure 3.11: Regional CO_2 surface fluxes from April 2009 to April 2010. CarbonTracker noreject follows CarbonTracker baseline outside South America. The uncertainties shown for Carbontracker are aggregated spatially but not temporally. See figure 3.9 for details.

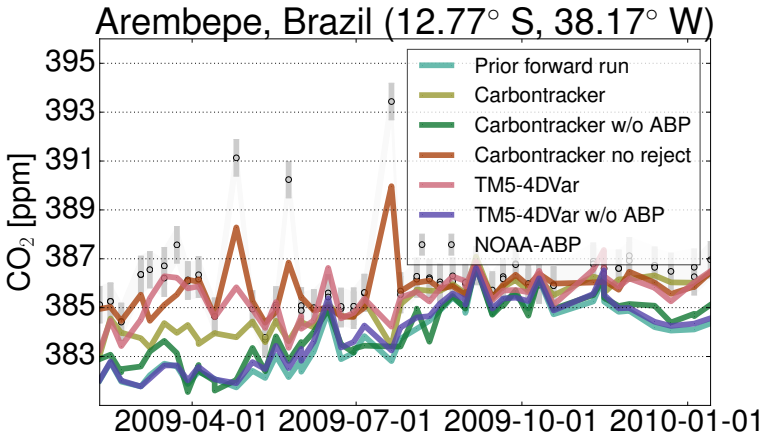


Figure 3.12: Time series of CO_2 concentration in Arembepe, Brazil at the east coast of South America. The two "without ABP" runs show the concentrations when the models do not assimilate data from the Arembepe site. CarbonTracker noreject shows the concentrations for CarbonTracker with disabled outlier rejection. The time series ends after January 2010, because data at Arembepe is only available in obspack PROTOTYPE v1.0.2 2013-01-28 from NOAA Environmental Sciences Division, Oak Ridge National Laboratory (2013) until then.

TM5-4DVar's flux anomaly in South America The time series of South American surface fluxes in Fig. 3.11 reveals that the flux differences in South America stem from particularly large emission estimates in summer 2009 by TM5-4DVar. The temporal structure of TM5-4DVar fluxes for the Indian Ocean as well as the Pacific Ocean, suggest that ocean uptake compensates for the large South America source to match the hemispheric flux budget.

South America suffers from sparseness of observational constraints such that validation of the estimated surface fluxes via comparison of measured and modelled atmospheric CO_2 concentrations is difficult. Aircraft measurements regularly conducted in South America do not provide deeper insight, because they have a data gap in the critical time between June and August 2009. The only other site that is close to the South America flux region is Arembepe in Brazil (ABP, 12.77° S, 38.17° W), a ground sampling station which is used as constraint within our data assimilation exercise.

To check its impact on the fluxes, we perform a sensitivity run which does not assimilate Arembepe. In this run both models are similarly good at matching modelled a posteriori and measured CO_2 concentrations in Arembepe and mostly follow the prior (see Fig. 3.12). When assimilating observations from Arembepe however, TM5-4DVar closely follows the observations in spring 2009 while CarbonTracker only moves half-ways from the prior to the observations. This can be explained by the outlier-rejection in CarbonTracker: when the difference between the model and a measurement is more than three times the estimated representativeness error of the measurement, CarbonTracker ignores the measurement as outlier. As marine boundary layer site, Arembepe is assigned a representativeness error of only 0.75 ppm, so CarbonTracker ignores most measurements before May 2009.

The aggregated fluxes in Fig. 3.10 show that assimilating the measurements in Arembepe has a significant effect on the a posteriori fluxes of TM5-4DVar. When taking out Arembepe from the baseline run, TM5-4DVar's attribution of fluxes shifts: the sinks in the Pacific and the Indian Ocean weaken while the strong source in South America disappears. The time series in Fig.

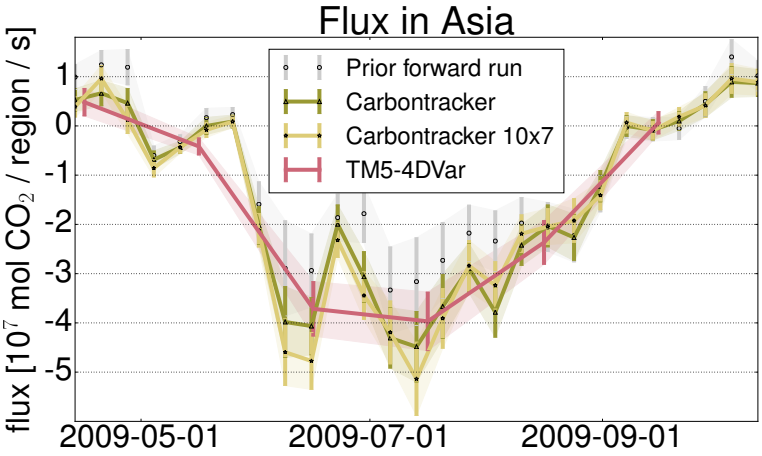


Figure 3.13: CO_2 surface fluxes during summer 2009 in Asia. The Prior Forward Run shows the prior fluxes aggregated to the bin-size of the weekly Carbontracker scaling factors.

3.11 provide a temporal fingerprint of the flux difference due to removing Arembepe from the assimilation which identifies the changes in the Pacific and the Indian Ocean as compensation for the removal of the strong source in South America.

The flux changes in CarbonTracker with assimilating Arembepe are within the estimated uncertainties, in the yearly aggregated fluxes as well as in the time series. Disabling the outlier rejection in CarbonTracker causes the modelled concentrations to follow the observations much more closely. Carbontracker specifies the flux uncertainty relative to the total flux, which in April and May 2009 yields a lower uncertainty than that from TM5-4DVar, which can cause the flux to change less than in TM5-4DVar in those months, leading to the strong reaction of the outlier rejection. But as shown in Fig. 3.11 neither Carbontracker with nor without the outlier rejection shows the additional source seen in TM5-4DVar between July and August 2009, where the flux uncertainty of both models differs by less than 10%. Also it does not show the compensation fluxes TM5-4DVar gives in the oceans. So we can reject the theory that these differences are caused by the outlier rejection or inconsistencies in the prior flux uncertainty.

The fluxes induced by assimilating Arembepe show that TM5-4DVar is more susceptible than CarbonTracker to the effect of single measurement sites in regions with very low observation density.

CarbonTracker with longer assimilation window Figure 3.10 shows that when increasing the assimilation time window of CarbonTracker to 5×20 days ("5×20"), CarbonTracker yields roughly the same aggregated flux for Asia as TM5-4DVar.

The time series in Fig. 3.13 suggests that the change in the CarbonTracker estimate of Asian fluxes when going to the longer assimilation window originates from high frequency adjustments to the prior fluxes. If the biosphere model needs to be corrected for only one week, the run with weekly flux bins can adjust that week separately while the run with 20 day flux bins has to adjust a full 20 day period. To test this theory, we verified that a run with an assimilation window consisting of ten one-week cycles yields a similar Asian sink as the run with five one-week cycles (1.84 instead of 1.61 PgC a^{-1}) which does not increase further when going to fifteen one-week

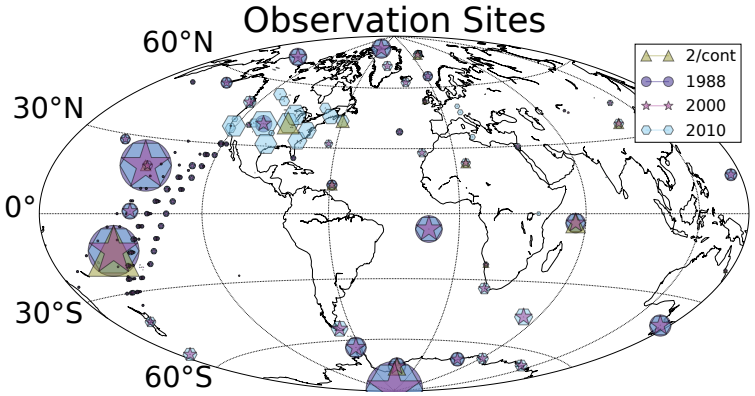


Figure 3.14: Observation sites assimilated by the '2/cont', '1988', '2000' and '2010' historical runs. The size of the symbol is proportional to $\frac{\sqrt{N}}{\sigma}$ with N the number of available measurements and σ their uncertainty estimate. The larger the symbol, the larger the weight of the site in the model.

cycles (not shown), while a run with three 20 day cycles yields a similar Asian sink as the run with five 20 day cycles (2.22 instead of 2.25 PgC a^{-1}).

For a quantitative discussion of the propagation of aggregation errors see Turner and Jacob (2015). Our findings suggest that there is an impact of roughly 0.5 PgC a^{-1} from high frequency mismatches between the prior model and the measured concentrations during the Asian summer which cannot be corrected accurately with a bin-size of 20 days or more.

In summary we see good agreement for the baseline fluxes between CarbonTracker and TM5-4DVar on a global scale and for most continents and oceans. The mismatch of the fluxes in South America, the Indian Ocean and Asia can be traced back to two distinct effects: a different flux response in regions with very limited observation coverage and using weekly (CarbonTracker) or monthly (TM5-4DVar) adjustments to account for mismatches on shorter time scales.

3.2.2.2 Sensitivity to observation coverage

In order to assess the importance of data density and coverage on the two DA systems, we follow the approach which Bruhwiler et al. (2011) used to analyse the performance of their initial version of a fixed-lag Ensemble Kalman Smoother (Bruhwiler et al., 2005). We carry out 5 "historical" model-runs where we stepwise increase the number of assimilated observation sites, mostly following the historical availability of data. The first run, termed "2/cont", assimilates observations from up to 2 sites per continent. It represents an extremely sparse observation network with different sampling frequencies per site. The runs "1988" and "2000" assimilate observations from all sites that were active in the years 1988 and 2000, respectively. The "2000" run assimilates roughly the same number of observations as our baseline run. The run "2010" uses all sites which were active in the year 2010 except for Arembepe. We exclude Arembepe from the "2010" run, because as shown in the section 3.2.2.1 the different treatment of the observations there would dominate the flux changes and as such mask other effects. Figure 3.14 illustrates the observation density and coverage for the different historical runs while Table A.1 lists the sites included for all the historical runs.

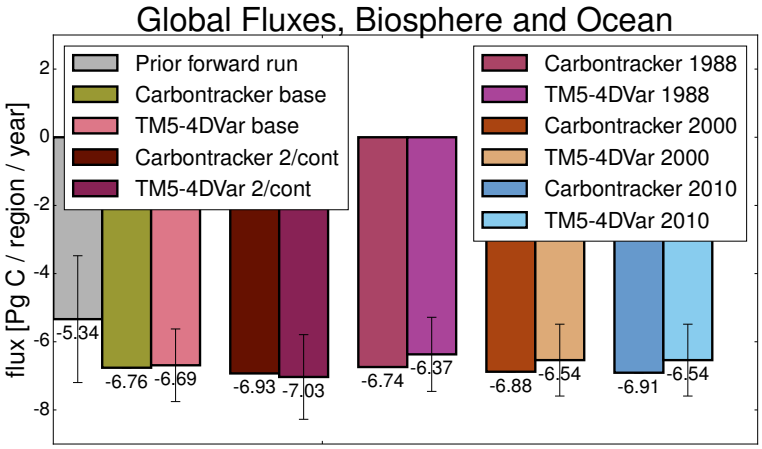


Figure 3.15: Globally aggregated surface fluxes estimated by the model runs indicated in the legend. In all aggregated flux bar charts the uncertainties are estimated by TM5-4DVar.

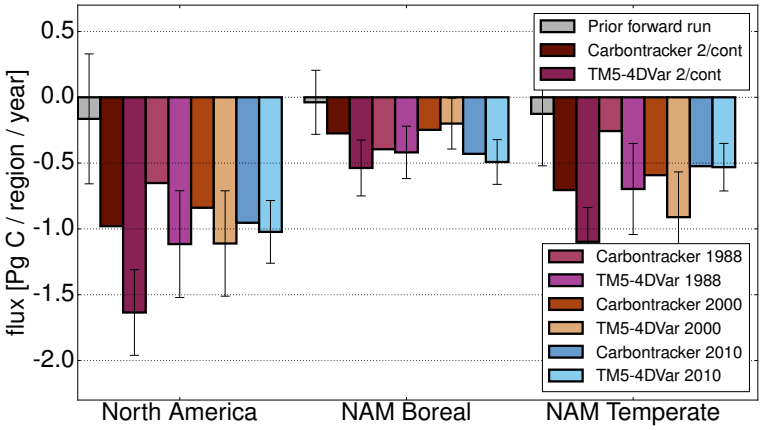


Figure 3.16: Fluxes for CarbonTracker and TM5-4DVar from April 2009 to April 2010 separated into the two Transcom Regions in North America.

Figure 3.15 shows the globally aggregated prior and a posteriori fluxes for the baseline setup and each of the historical runs. All the historical runs for both models, CarbonTracker as well as TM5-4DVar, yield consistent estimates of the global (biospheric and oceanic) carbon sink. The results differ by a few tenths of a PgC a^{-1} which is well below the TM5-4DVar uncertainty estimate of about 1 PgC a^{-1}. This consistency is expected since the global carbon sink is well constrained by the trend in global background concentrations. Compared to the prior, all runs indicate a stronger sink by more than 1 PgC a^{-1}. The global flux estimate is robust against changes in the observation coverage and against the choice of the inverse method. Global scale fluxes are also consistent with the 2013B estimates from CarbonTracker North America (NOAA, ESRL).[3] NOAA shows a global sink of 6.79 ± 6.86 PgC for 2009 while we see values between 6.37 and 7.03 PgC for April 2009 to April 2010, depending on the observation data we assimilate.

[3]The 2013B release of the estimated fluxes of Carbontracker North America (NOAA, ESRL) is available from esrl.noaa.gov/gmd/ccgg/carbontracker/CT2013B/fluxtimeseries.php?region=Global.

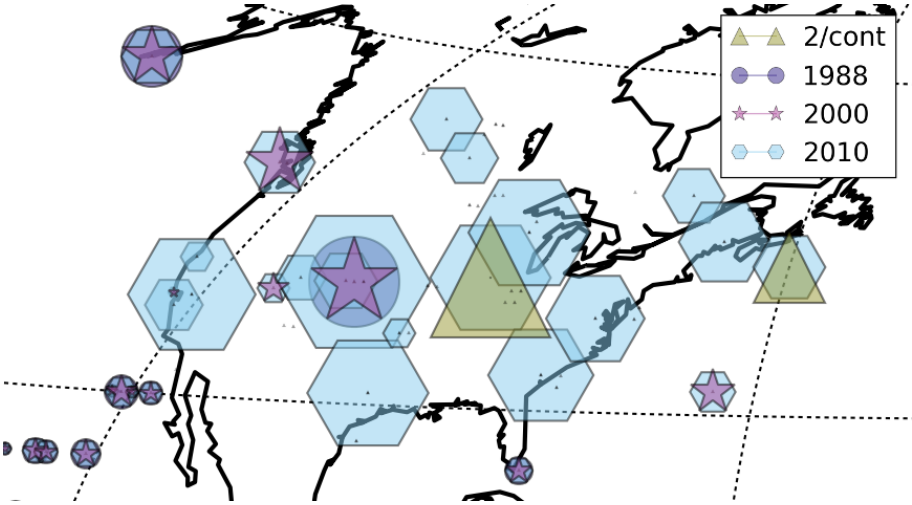


Figure 3.17: Visualization of the weight of the measurement sites which are assimilated in North America in the respective runs.

On the continental scale we take a closer look at North America, since changes in observation density are historically most pronounced there. Figure 3.16 shows that TM5-4DVar and CarbonTracker fluxes for North America become more similar the denser the observation network becomes, with almost the same flux estimate in the "2010" setup in which the DA systems assimilate more than 15 sites on the North American continent (see Fig. 3.17). This good match of both methods suggests that the density of observation sites in North America suffices to optimize continental scale fluxes with some degree of certainty. Separating the fluxes of the two North America Transcom regions (Fig. 3.16) shows that for the more homogeneous transcom region in Boreal North America the results from both methods are already converged with the observation coverage in the "1988" run, while in the more heterogeneous North American Temperate region with much agricultural activity, the methods only converge in the "2010" setup.

The stronger land sink seen by TM5-4DVar for "2/cont" stems from assimilating only two sites: a site in West Branch in Iowa, USA (WBI, 41.7° N, 91.4° W), in the US corn belt, and a site on Sable Islands, Nova Scotia, Canada (WSA, 43,9° N, 60.0° W). In TM5-4DVar, the strong summer sink near West Branch dominates the North America fluxes and increases the sink from roughly 1 PgC a^{-1} in the "2010" run to more than 1.6 PgC a^{-1} in the "2/cont" run. CarbonTracker is less susceptible to this effect than TM5-4DVar, because its ecoregion approach enforces a correlation between the fluxes for all regions in the corn belt as well as for all regions with grassland – both region-types span the area from the southern parts of North America up to the border of Canada. This makes it more likely that a potential flux adjustment is constrained by more than one site which gives CarbonTracker a stronger meridional coupling. Since meridional mixing is much slower than zonal mixing, stronger meridional coupling forces a larger region to change in the same way. For example adjusting the flux in the corn belt yields concentration changes all over North America (downwind of the corn belt ecoregion). This effect cannot be replicated with a single global correlation length parameter, because there are other areas where this would be unrealistic, for example the European continent with its very heterogeneous land use.

On the other hand, the overall North American sink of 0.65 PgC a^{-1} estimated by CarbonTracker in the "1988" run is 30% lower than the sink of 0.95 PgC a^{-1} in the "2010" run, while in

TM5-4DVar the "1988" and the "2010" run differ only by 10% (0.1 PgC a^{-1}). The difference between the results for the "2000" and the "2010" runs in North America is on the order of 0.1 PgC a^{-1} for both models, but in different directions. So with low observation coverage, the quality of the inversion in either system depends on the exact distribution of the observations. This suggests that with the coverage from "2000", we need to assume a minimum uncertainty of 0.25 PgC a^{-1} from only the choice of the inverse method. For the coverage from "2010" this is down to less than 0.1 PgC a^{-1} in North America.

The strong reduction of the uncertainty estimate in the North America fluxes of TM5-4DVar in the "2/cont" run, despite assimilating only 2 sites in North America, shows the sensitivity of these uncertainty estimates to the raw number of assimilated observations. It proves that the actual structure of the observational network has to be taken into account when interpreting the reduction of model-estimated uncertainty.

Overall our results show that the current observation coverage in North America allows estimating robust fluxes on continental scales and on the scales of transcom regions (as defined by Gurney et al., 2000). The historically improving agreement between both models for the aggregated North American fluxes and the two transcom regions in North America suggests that increasing the observation coverage allows getting robust fluxes on even smaller scales. To get better fluxes, we have to assimilate more observations with different coverage to improve the overall observation density.

3.3 Conclusions

Our study evaluates the performance of the data assimilation models CarbonTracker and TM5-4DVar by comparing their a posteriori CO_2 concentration fields to measurements and by comparing their a posteriori surface fluxes. We test the sensitivity of the a posteriori CO_2 fluxes to model parameters and data coverage. To analyse the impact of the inverse method and the flux representation, the models run in setups which are close to their production settings but use harmonized input data, tracer transport model, prior flux and prior flux covariance estimates. A caveat applies since prior fluxes and prior flux uncertainties cannot be made identical due to differences in how the state vectors of the two methods are setup: Carbontracker optimizes weekly ecosystem-wide fluxes while TM5-4DVar optimizes monthly fluxes on a regular longitude-latitude grid.

Both inverse models yield CO_2 concentration fields of comparable quality. We show that increasing the length of the assimilation time window of CarbonTracker to five bins of twenty days or ten bins of seven days gives a good agreement to observations in Antarctica which are underestimated in summer when using the default setup with an assimilation window of only five weeks. With these longer windows, the difference of the bias of the models at non-assimilated measurement sites is lower than the uncertainty of the bias due to the limited number of non-assimilated sites. This has two implications: first, the differences between the a posteriori fluxes provide a lower estimate of the uncertainty due to the choice of the optimization method, and second, a choice between the two systems may reduce to practical considerations, such as (a) Carbontracker is easily parallelisable because of the ensemble structure, but (b) TM5-4DVar yields defined uncertainties over long time flux integrals which have to be approximated in Carbontracker, or (c) TM5-4DVar requires an adjoint of the transport model, Carbontracker does not.

The a posteriori fluxes from both models are in good agreement on a global scale, but on continental scales they show significant differences, most noticeably in South America which has very sparse coverage of observation sites. Investigating the flux time series allows tracing

these differences back to spurious flux adjustments in TM5-4DVar for South America due to assimilating observations from a single site in Arembepe, Brazil, along with compensating fluxes in the oceans. Also we see a difference in the adjustment of Asian fluxes, but an additional CarbonTracker run with a coarser temporal flux adjustment bin size of 20 days gives similar fluxes in Asia as TM5-4DVar. Here, the flux time series reveal that part of the weaker sink in CarbonTracker with smaller bin size stems from high frequency changes which cannot be represented with the monthly binning of flux-adaptation in TM5-4DVar and the CarbonTracker run with bins of 20 days. The impact of this effect on the fluxes in Asia is 0.5 PgC a^{-1}.

To better analyse the sensitivity of both models to the observation coverage, we run the models with collections of measurement sites selected by historical availability. In North America, where the change of observation coverage is most pronounced, fluxes estimated with the observation network from 2000 differ by 0.25 PgC a^{-1}. With the measurement network from 2010, the difference reduces to 0.1 PgC a^{-1}. In other regions the coverage from 2010 is closest to the coverage in North America from 2000, so 0.25 PgC a^{-1} can serve as lower limit for the uncertainty due to changing the method

TM5-4DVar has a stronger response to the data coverage than CarbonTracker. This shows that the ecoregion approach in CarbonTracker with its stronger meridional coupling of fluxes and observations makes CarbonTracker less susceptible to changes in the distribution and density of observations than the simple global flux covariance in TM5-4DVar. As such it might be useful to reuse CarbonTracker's spatial flux correlation structure in TM5-4DVar.

The improved agreement between both models when adding observation sites indicates that the coverage of observation sites in North America should be sufficient to yield robust fluxes on a continental scale when only considering the uncertainty from the inverse methods and the flux representation.

In layman's terms, this study showed that with small parameter adjustments the two analysed systems for estimating surface fluxes (CarbonTracker and TM5-4DVar) yield concentration fields of similar quality, quantified by the mismatch to sites they did not use for the optimization of the fluxes.

Adding to the conclusion of Babenhauserheide et al. (2015), this study showed that the main task for getting better flux estimates is assimilating more measurements which complement the existing network.

I chose to continue this study with CarbonTracker, because it should be more robust against changing observation coverage, which is important for assimilating satellite data, it is easy to parallelize in case an application requires better performance, and because due to requiring no adjoint it will be much easier to connect to the other transport models which are developed at our institute (IMK-ASF), including the ICON-ART system which builds on ICON, the new multi-scale meteorology model from the German weather service (dwd).

With this choice, I started with another vector to understand uncertainty: Assimilating remote sensing total column measurements from the TCCON network which provide constraints for the free troposphere. This allows estimating uncertainties with an independent validation which has a different sensitivity than the measurements used in this part of the study and which can reduce the uncertainty of fluxes when used as a constraint in the inverse model. The structure of TCCON measurements is similar to the existing in-situ measurements, with repeated measurements at fixed locations, so they are well suited as a first step into assimilating column measurements.

Chapter 4

The added value of total column observations for CO_2 inverse modelling

After showing in the previous chapter that the main requirement for more robust flux estimates are additional measurements, this part of the study investigates joint assimilation of total column measurements from the TCCON network alongside in-situ measurements. Remote sensing techniques measure the CO_2 concentration not only close to the ground but also in the free troposphere and higher (see section 2.2.2). They complement in-situ measurements, because they include more information from spatially remote flux regions.

4.1 Observations and Methods

The total column data from TCCON, described in section 2.2.2, provides the currently best measurements of the aggregated CO_2 column. The average station-by-station bias is less than 0.2 ppm (Wunch et al., 2011a). This study uses the release GGG2014 from the TCCON website.[1] In addition to CO_2 total columns the release provides daily a-priori data and averaging kernels binned by solar zenith angle for all sites, which makes it easy to assimilate the measurements.

The in-situ measurements assimilated here are taken from the obspack collection (NOAA Environmental Sciences Division, Oak Ridge National Laboratory, 2013), a collection of high-precision measurements from ground stations and aircrafts described in section 2.2.1 and also used in chapter 3.

Every 30 minutes of model-time the employed transport model (Transport Model 5, TM5, see section 2.4.2) calculates a CO_2 vertical concentration profile of CO_2 at each TCCON site. Carbontracker represents model covariance via an ensemble of runs with disturbed fluxes. To assimilate data from TCCON CarbonTracker calculates the modelled total column for each ensemble member individually and use the ensemble spread as the model uncertainty.

Comparison of the modelled profile with a total column measurement requires simulating a measurement in the model, taking the measurement sensitivity into account (Basu et al., 2011). To simulate the measurement, the column averaging kernel of the TCCON measurement ($\vec{A}_c$) as defined in equation (2.2) and the prior of the measurement $\vec{x}_b$ are interpolated to the mean values of the model layers.[2] Then the model values $\vec{x}$ and the interpolated prior $\vec{x}'_b$ are aggregated,

[1] TCCON data is available from the website tccon.ornl.gov

[2] Linear interpolation between pressure levels incurs some linearization errors. These could be reduced in a future version by weighting the contribution of the different layers to the interpolated prior concentrations and

weighted by the interpolated column averaging kernel ($\vec{A}'_c$) to get the total fraction of CO_2 in the column.

$$CO_2 = \sum_i \left(x'_{b,i}\right) + \sum_i \left(A'_{c,i}(x_i - x'_{b,i})\right) \tag{4.1}$$

with the elements of the apriori profile $x'_{b,i}$, the model profile x_i, and the column averaging kernel $A_{c,i}$ representing layer means. A detailed explanation of the column averaging process used is given in the auxiliary information on the TCCON website.[3]

4.2 Viability of assimilating total column data

Before assimilating total column measurements, it is necessary to get an estimate of the performance of the model's vertical transport. Model transport can be evaluated by comparing modelled vertical concentration profiles of CO_2 to measured ones. The assimilation can only provide useful information about remote regions if vertical concentration profiles of CO_2 exported from the model are sufficiently close to real profiles. Strong mismatches in a large fraction of the profiles would indicate problems in vertical transport in the model: The region of highest influence on the profile would differ from real influences. This would cause additional misattribution of fluxes, making total column assimilation infeasible. And to give a robust assimilation, the shape of the profiles must change more slowly with time than the shape of real measurements, to ensure that adaptions due to model artefacts coinciding with measurements are much smaller than adaptions from measured changes in the atmosphere. Export of the profiles from the transport model is realized using a module ported from TM5-4DVar (Basu et al., 2013) to Carbontracker.

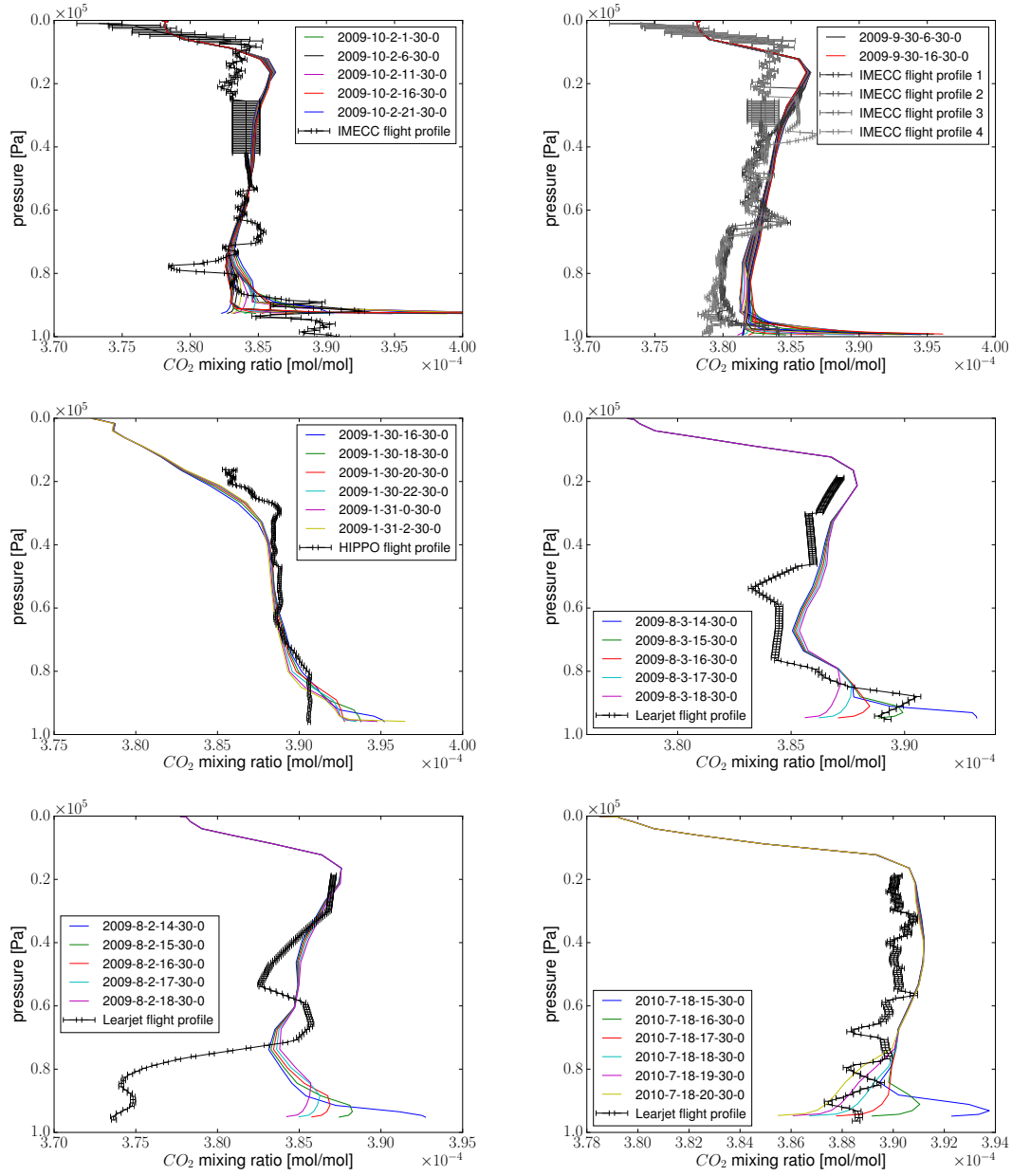
Observations to test the vertical transport are available from aircraft profiles measurements conducted above TCCON stations. Figure 4.1 compares profiles exported from CarbonTracker with IMECC flight profiles (Feist et al., 2010; Messerschmidt et al., 2011) over Karlsruhe and over Bialystok, with flight profiles over Lamont from Learjet on consecutive days in August 2009 (see Appendix A.3), from a HIPPO flight on 30 January 2009 and from Learjet flights on 18 July 2010 and 31 July 2009. The shown Carbontracker profiles were optimized using the setup described in chapter 3.

To make assimilation of total column measurements viable, most modelled profiles need to be similar to measured profiles without systematic bias by height. Such a systematic bias would lead to stray adaptions of the profiles because the adaptions can only take the aggregated column into account, with some additional information where signals from different stations overlap. Also the match must not be better than the precision of the total column measurements.

Five of the six flights show agreement between measured profile and model profile in the shape of the profile. They show that CarbonTracker can reproduce the measurements for profiles whose concentrations change by more than 2% with height. However, the comparison over Lamont on 2 August 2009 (Figure 4.1a) shows a mismatch of 10 ppm in the lowest 2km of the atmosphere (1000hPa to 800hPa). Aggregated over the total column, this would still be a mismatch of roughly

column averaging kernel with the pressure given by TCCON for the layers. Due to the low variability of the prior profiles and averaging kernel, the aggregation ignores this effect.

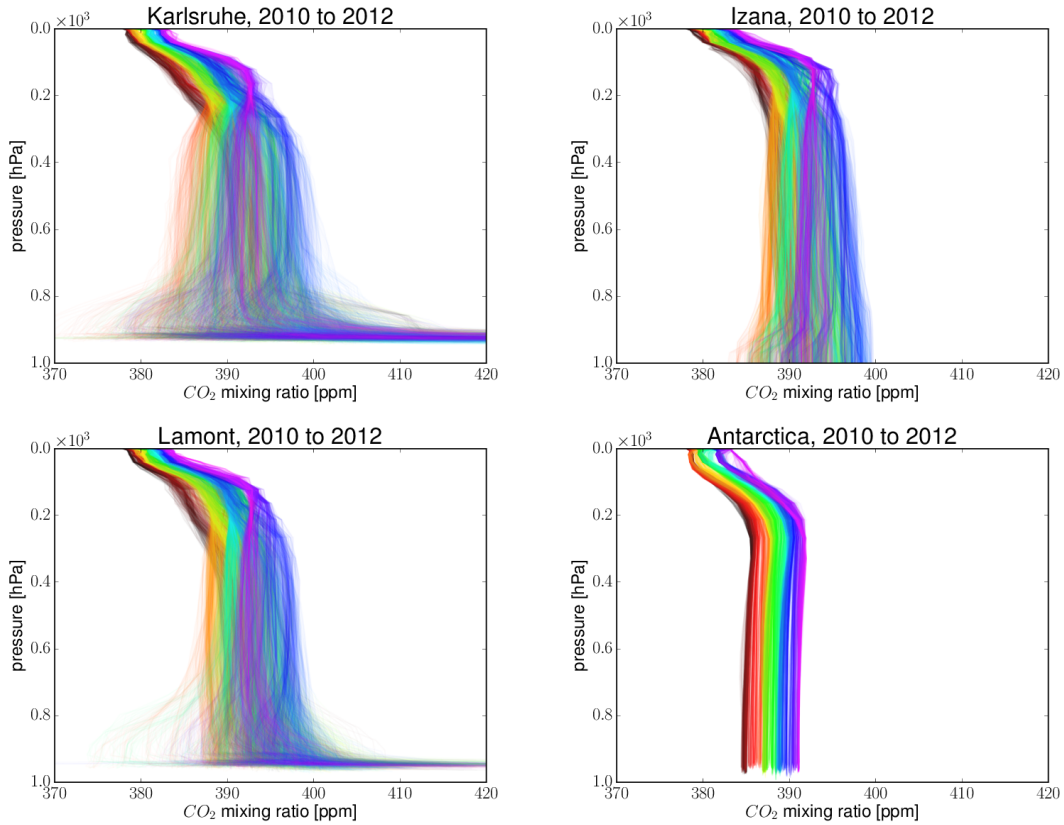
[3] The auxilliary information on the TCCON website gives detailed instructions for comparing total column measurements with models along with estimates of the resulting error size when a given step is omitted. The data on the auxilliary information website (averaging kernel and priors) is outdated, though, since more complete data is included directly in the NetCDF4 based GGG2014 data product. https://tccon-wiki.caltech.edu/Network_Policy/Data_Use_Policy/Auxiliary_Data



(a) IMECC Karlsruhe, 2009-10-02 (top)
HIPPO Lamont 2009-01-30 (middle),
LEARJET Lamont 2009-08-02 (bottom)

(b) IMECC Bialystok, 2009-09-30 (top)
LEARJET Lamont 2009-08-03 (middle),
LEARJET Lamont 2010-07-18 (bottom)

Figure 4.1: Comparison of modelled profiles with flight profiles from the IMECC campaigns over Karlsruhe and Bialystok (Feist et al., 2010; Messerschmidt et al., 2011) and the HIPPO (Wofsy, 2011) and LEARJET (A.3) campaigns over Lamont. Differences in the height of the lower measurements originate in differences between the model pressure and the measured pressure. The information of the specific flights is given in the sub-captions 4.1a and 4.1b.



(a) Karlsruhe (top), Lamont (bottom)

(b) Izana (top), Antarctica (bottom)

Figure 4.2: Modelled CO_2 profiles assimilating only in-situ data from ground-sites and aircraft missions. They are coloured by time, from red 2010 to blue 2012.

2 ppm, clearly detectable with the precision of TCCON measurements of less than 1 ppm (see table 2.1).

Correcting a profile like the one from Lamont on 2 August 2009 (Figure 4.1a) using only total column measurements could lead to a spurious scaling of the full profile. When combined with measurements close to the surface, though, such a total column measurement adds the information that the fluxes to adjust need to be nearby. An incorrect adjustment of remote fluxes would change concentrations higher in the atmosphere and as such lead to a smaller adjustment of the profile close to the surface. This would lead to a mismatch at in-situ sites, thus matching both total column and in-situ measurements provides information about the magnitude of the required concentration change and about the distance of the assimilated measurements from the region where the flux needs to be adjusted.

Analysing the long term evolution of model profiles allows to verify, that these results are robust. To this end, Figure 4.2 shows the development of the shapes of profiles.

Since there are mostly smooth variation in the profiles above 800hPa (2km), optimizing them with total column data should work well. The strong changes below 800hPa suggest, that assimilating both total column data and in-situ ground data should provide sufficient information

to discriminate between local and remote effects from most profiles, not just those validated by aircraft flights.

Together these checks provide the confidence that assimilating total column data can provide meaningful additional constraints for surface flux estimation.

Since chapter 3 showed that the impact from the choice of model becomes less important with higher measurement density, CarbonTracker was chosen for a practical and strategic reason: It has low requirements on the transport model. This makes it easier to implement new types of measurements and to combine it with other models developed at KIT IMK-ASF.

The changes to the setup from chapter 3 are:

- it uses a gridded version of CarbonTracker Europe, as described in section 2.5.3,
- the transport model 5 (TM5) was adapted to export profiles,
- the runs use only 30 ensemble members instead of the 300 used in chapter 3. This increases linearisation errors because the ensemble cannot capture the exact distribution but decreases the runtime of the comparison. A result from this limitation is higher scatter of the modelled concentrations,
- assimilation of total column measurements is realized by chaining the measured total column values into the regular observation vector with the uncertainty calculated from the number of measurements at that site on that day.

In the following sections total column measurements are assimilated. First they are assimilated as the only constraint to estimate their representativeness error (section 4.3), then they are used in joint inversions with in-situ measurements (section 4.4).

4.3 Estimating the representativeness error

As discussed in chapter 3, the quality of flux estimates depends on the choice of the representativeness error for the assimilated concentrations. If the representativeness error estimate is too large, the uncertainty of the fluxes is higher than necessary and the flux sticks too much to the prior. If the representativeness error is too small, the model becomes unstable and fails to reproduce validation measurements. A suitable representativeness error needs to balance information content against unphysically strong adaptation to single sites.

In this section a suitable representativeness error is estimated by varying the ascribed uncertainty for TCCON measurements. An ideal choice of the representativeness error minimizes the mismatch of the modelled concentration fields to in-situ measurements, along with minimizing the estimated uncertainty of fluxes and keeping the estimated fluxes plausible.

Ascribed uncertainty follows the definition of the representativeness error as in chapter 3. Following Chevallier et al. (2011) the representativeness error is given per day. This avoids inter-day biases from different weighting due to variable measurement counts.

$$\sigma = \sigma_{\text{day}} \cdot \sqrt{N} \tag{4.2}$$

with σ the uncertainty for a single measurement, N the number of measurements on the given day and σ_{day} the daily uncertainty.

Figure 4.3: Validation of TCCON assimilation with varying representativeness error using the yearly mean flux on 1°×1° latitude×longitude grid, 2010-12-03 to 2011-12-02. The flux maps show the variation of the daily uncertainty from 2 ppm in the top right histogram, reduced in a counter-clockwise circle over 0.5 ppm in the top left – the choice of Chevallier et al. (2011) – and 0.34 ppm in the bottom left and 0.2 ppm in bottom centre – the standard choice in this study – to 0.08 ppm in the bottom right.

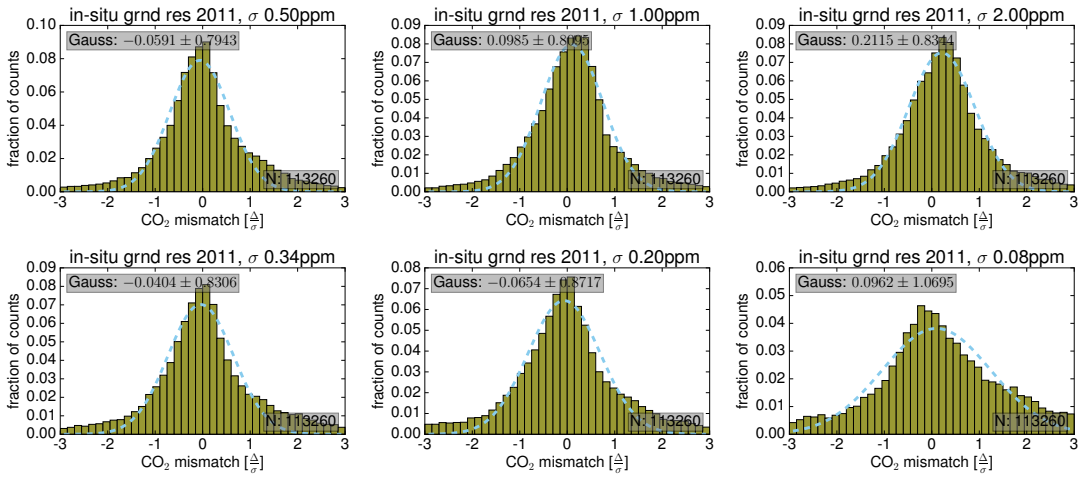
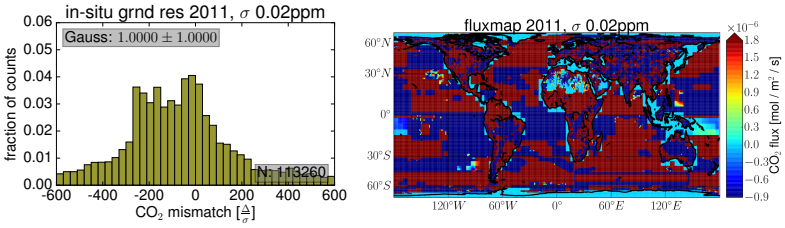


Figure 4.4: Validation of TCCON assimilation with varying representativeness error using the mismatch against the non-assimilated in-situ measurements, 2010-12-03 to 2011-12-02. The histograms show the variation of the daily uncertainty from 2 ppm in the top right histogram, reduced in a counter-clockwise circle over 0.5 ppm in the top left – the choice of Chevallier et al. (2011) – and 0.34 ppm in the bottom left and 0.2 ppm in bottom centre – the standard choice in this study – to 0.08 ppm in the bottom right.

Chevallier et al. (2011) choose a daily uncertainty of 0.5 ppm, doubled in Europe where the density of TCCON sites is highest. This value can be improved by varying the representativeness error in model runs which only assimilate TCCON measurements and comparing the resulting CO_2 concentration fields against measurements from the in-situ network. To avoid potential biases from the limited set of validation sites (demonstrated in section 4.4.1), the optimized CO_2 concentration fields are compared against all in-situ measurements.

Figure 4.4 compares different choices for the daily uncertainty for the time frame from December 2010 to December 2011. The uncertainty of the data in the histograms varies in a counter-clockwise circle, from 2.0 ppm in the top right to 0.08 ppm in the bottom right. This contrasts the two extremes of the values. The mean yearly fluxes on a $1°\times1°$ scale are shown in Figure 4.3. For 0.34 ppm and 0.2 ppm the histograms show the best fit to ground measurements. The flux maps show the strengthening patterns of flux changes which at 0.08 ppm uncertainty contain strong dipole patterns, forming a red-blue source-sink pattern throughout Eurasia and North America. Figure 4.5 shows that the model behaviour becomes unphysical with a daily uncertainty of only 0.02 ppm. The histogram shown in Figure 4.5 shows strong distortions. Its shape is similar to the shape of histogram for 0.08 ppm daily uncertainty, though more extreme: the tip moves towards underestimation of the measurements. The flux map for 0.02 ppm in Figure 4.5b shows that these distortions originate in strong oscillations in the flux patterns. All runs shown in these histograms start in 2010 with an initial concentration field optimized for 2009, therefore they require an initial global source to increase the CO_2 concentrations by 2 ppm.

Unphysical behaviour can be detected by analysing the regionally aggregated fluxes for plausibility. Fig 4.6 shows the fluxes corresponding to the histograms shown in Figure 4.4. The fluxes are within the 1σ uncertainty of the in-situ fluxes for choices of the daily representativeness error higher than 0.1 ppm. The validity of the yearly uncertainty estimates of the in-situ flux estimates



(a) 2010-12-03 to 2011-12-02

(b) 2010-12-03 to 2011-12-02

Figure 4.5: Results with a daily uncertainty for TCCON sites of only 0.02 ppm. Mismatch between the modelled concentration field and in-situ sites, modelled value minus measurement, and a 1°×1° world map of fluxes. The spread of more than the absolute value of CO_2 and the pattern of neighbouring strong sources and sinks in Fig 4.5b show the effect of oscillations and unphysical behaviour due to over-fitting measurement noise as signal.

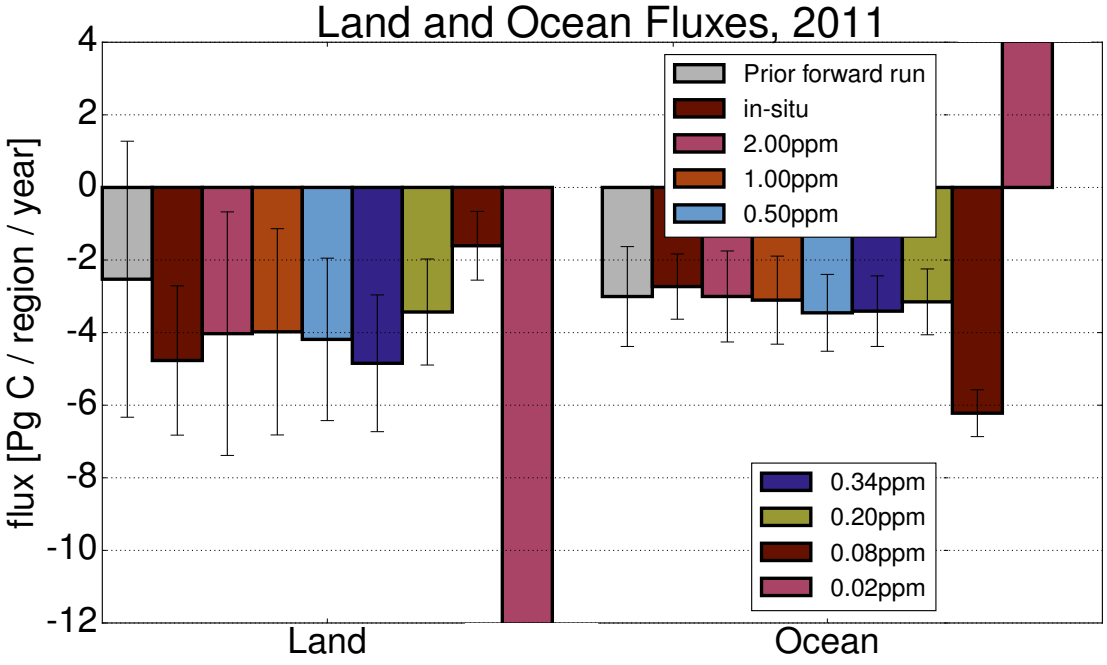


Figure 4.6: Global flux estimates for 2011 (from 2010-12-03 to 2011-12-02) using different values for the representativeness error. The runs start in 2010 with the concentration field of 2009, which starts about 2 ppm too low, due to the annual increase of the CO_2 concentration by about 2 ppm.

is shown in Basu et al. (2011) and Babenhauserheide et al. (2015), so fluxes outside this range are an indicator of unphysical behaviour.

Combining the bias against in-situ sites and the flux uncertainty estimates provides an objective measure to choose the uncertainty. This measure is shown in Figures 4.7a and 4.7b for land and ocean fluxes, respectively. The figures show a C-shape, since the prior fluxes are too low by roughly 1 PgC/a (as verified in chapter 3), so a too high representativeness error sticks too strongly to the prior which results in a too weak sink and as such too quickly rising CO_2 concentrations. The unphysical values for 0.02 ppm are left out because they lie outside the scales of the graph.

These analyses show that using a representativeness error with daily uncertainty between 0.1 ppm and 0.5 ppm provides concentration fields consistent with in-situ measurements. For assimilation together with in-situ measurements, a value in the lower part of this range should be reasonable, because the in-situ measurements provide a second constraint to keep the concentration field physical. Using 0.2 ppm is on the order of the station-by-station bias of 0.2 ppm (Wunch et al., 2011b). This choice is reasonable because with an average wind speed of more than 4 ms^{-1}, one day of measurements samples a section of air with a length on the order of the length of the grid boxes of our transport model (500km). Due to this, the station-by-station bias limits the daily accuracy of the measurements.[4]

The remainder of this study uses a daily uncertainty of 0.2 ppm, about 40% of the value chosen by (Chevallier et al., 2011) but with some distance to the lowest reasonable values found here.

4.4 Joint assimilation of total column and in-situ measurements in Carbontracker

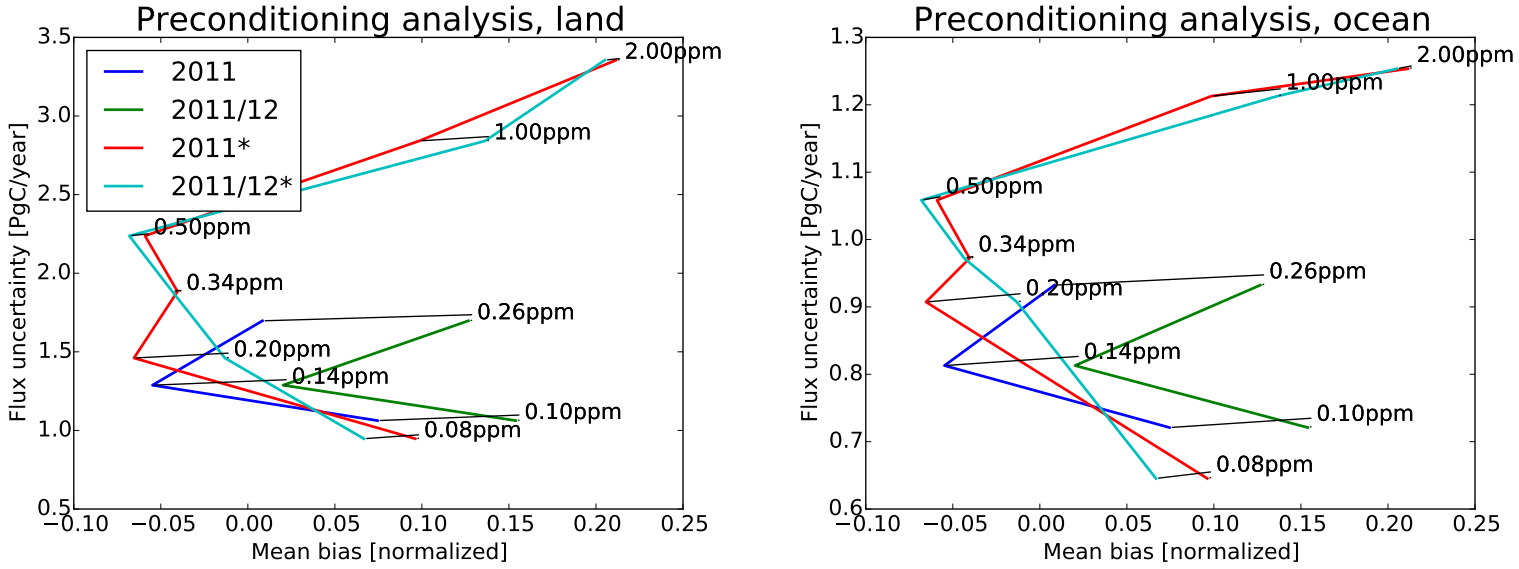
As shown in section 3, the main requirement for better flux estimmation is assimilating more measurements. In this section, assimilation of different combinations of measurement types is investigated (the combinations and shorthands used to reference them are shown in table 4.1). First the results of these assimilations are compared against observations to investigate their characteristics, then they are used to find a new estimate of Eurasian CO_2 fluxes und uncertainties.

4.4.1 Cross-comparison of modelled fields and sets of measurements

Figure 4.8 shows modelled CO_2 concentrations estimated by assimilating different sets of measurements compared with each of these sets of measurements. Additionally the concentration fields are compared against non-assimilated validation measurements (`V`). The different runs are specified using the shorthand defined in Table 4.1. Some aspects of Figure 4.8 will be investigated in more detail in section 4.4.2.

The cross-comparison uses a boxplot graph (McGill et al., 1978, Figure E). The black line in each box shows the median value of the mismatch between modelled and measured concentration for all the validation measurements. The coloured boxes contain 50% of the values closest to the

[4] The best choice changes with the grid-size of the transport model. With smaller grid boxes (i.e. 1°×1° longitude×latitude) the representativeness error becomes smaller, so that transport errors and genuine measurement errors become more important. With grid sizes of 1°×1° longitude×latitude for the flux representation and 6°×4° longitude×latitude for the transport used in this study, the representativeness error still dominates. Even with only 16 measurements per day, corresponding to 8 hours of measurement time with two measurements per hour, an uncertainty of 0.4 ppm per measurement due to non-systematic measurement noise (0.1% of the total concentration value and twice the station-by-station bias) would correspond to a daily uncertainty of 0.1 ppm, roughly half the uncertainty used here.



(a) Land

(b) Oceans

Figure 4.7: Analysis of the variation of the representativeness uncertainty. The graphs show the bias as estimated from the histograms for two different time periods. The runs marked by a * start in 2010 with an initial concentration which is too low by 2 ppm. Figure 4.6 shows the fluxes from these marked runs. The y-axis gives the flux uncertainty while the x-axis shows the normalized mean bias for the corresponding histograms with the labels marking the respective daily uncertainty. The upper arm of the C-shape stems from a too weak sink in the prior which leads to overestimation of the CO_2 concentration by the prior forward run, insufficiently corrected by the TCCON assimilation with a too high representativeness error. The lower arm stems from over-fitting the model. The runs with 0.02 ppm daily uncertainty are not shown because their values fall far outside the scales, with a bias of more than -100 (normalized).

median. The dashed lines with whiskers show the spread of 95% of the values, again counted from the median. The black circles outside the whiskers show the 2.5% highest and 2.5% lowest values. In addition to the median line, the red hexagons show the mean value of the mismatches. Finally the grey diamonds show the root mean square of the mismatches after removing all 2σ outliers to avoid giving a high weight to outliers.

The left part of the plot with grey background shows the comparison against in-situ validation data (`V`) as used in Babenhauserheide et al. (2015). The second part shows validation against ground-based in-situ measurements (`G`). The third part against aircraft-based in-situ measurements (`A`), the fourth against all TCCON sites in the GGG2014 R0 release (`Ta`) and the fourth against European TCCON sites (`Te`) except for Bremen and Paris which did not yet upload the data when this comparison was done.

Generally the different comparisons show a common pattern: The run with a priori fluxes overestimates the validation data by 2–4 ppm CO_2 on average (mean and median). This indicates underestimation of the carbon sink in the SiBCASA biosphere model. Assimilating any of the datasets reduces this mismatch to at most 1 ppm mean and median. Consequently, the RMSD excluding the 2σ outliers is between 3 and 4 ppm for the run with a priori fluxes while it is below 2 ppm for all comparisons except 3 sets: comparing assimilation of all TCCON sites and all in-situ sites against the ground sites (`TaGA_G`), comparing assimilation of all TCCON sites against the ground data (`Ta_G`) and comparing assimilation of all TCCON sites against aircraft measurements (`Ta_A`).

Aside from the different median and mean, there is a significant difference between comparison against validation data and the other runs. The 50% boxes of the comparison against validation hold only a fraction of the variability of the 50% boxes of other comparisons. This difference shows that the validation data (`V`) is not representative of the variability of the measurements, so a representative comparison of a TCCON assimilation must be validated against the full ground dataset (`G`), not against the limited validation dataset. This is most visible in the `Ta_V` and the `Ta_G` datasets. The modelled concentration fields which assimilated all TCCON sites show a mean mismatch against the validation measurements of more than 1 ppm while the mean mismatch

Table 4.1: Shorthand definition of runs with different assimilated measurements and validation measurements. A run is defined as `X_Y`, with `X` indicating the type of measurements used in the assimilation and `Y` indicating the type of measurement used for validation. The shorthands can be combined, for example as `GA_Te`: Assimilating in-situ ground (`G`) and in-situ aircraft (`A`) data and validating against European TCCON sites (`Te`).

shorthand	definition
`_`	No measurements are assimilated. In section 3 this is called the prior forward run.
`V`	Use only in-situ validation sites. This is only used do describe validation data.
`G`	Use in-situ measurements from ground. If this describes the assimilated data, the measurements exclude `V`, if it describes the validation data, the measurements include `V`.
`A`	In-situ measurements from aircraft.
`Te`	All TCCON sites within Europe.
`Tne`	All TCCON sites outside Europe.
`Ta`	All TCCON sites. `Ta` is equal to `TeTne`.

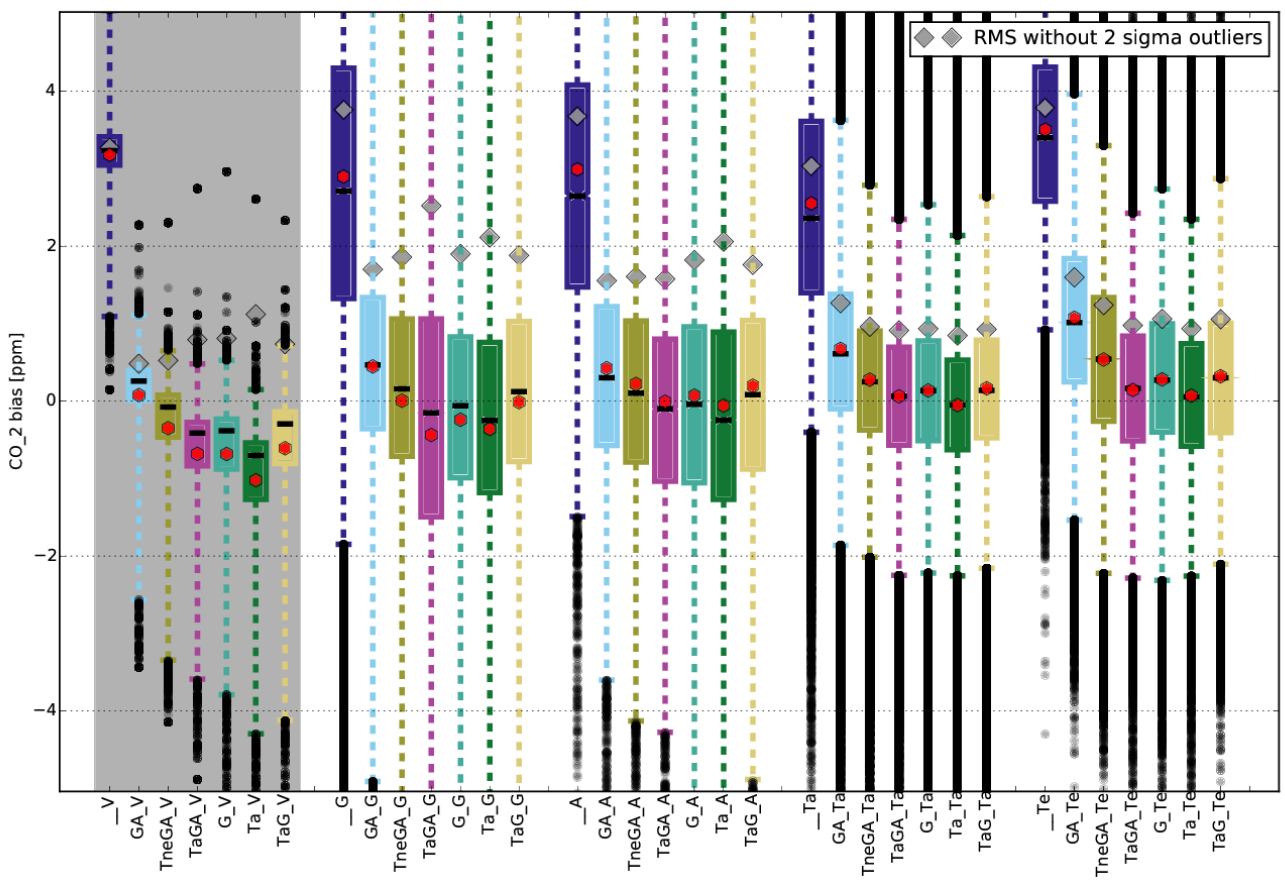


Figure 4.8: Cross-comparison of datasets using 3 years of data from 2009-04-03 to 2012-04-02. The x-axis shows the visualized data in the format defined in Table 4.1 as `X_Y`: assimilating `X` and comparing against `Y`. The different datasets are V, the validation sites from Babenhauserheide et al. (2015)[a], G, the ground-based in-situ sites, A, the aircraft-based in-situ measurements, Te, the European TCCON sites, Tne, the TCCON sites outside Europe and Ta, all TCCON sites. The black horizontal line is the median, the red hexagon the mean, and the grey diamond the root-mean-square of the values after taking out the 2σ outliers. The coloured bar contains 50% of the measurements, the dashes contain 95% of the data and the circles show outliers (the 2.5% highest and the 2.5% lowest datapoints).

[a]The validation includes the sites by site-code: BGU,CBA,CDL,DRP,EGB,FIK,FNS,HEI,JFJ,LJO,LPO,LUT,MEX,OTA,PDM,TOT,TRN

against all ground measurements is less than 0.5 ppm. The median shows a lower difference, but the median mismatch against the validation data is still two times as large as the the median mismatch against all ground data.

This has an important implication on the interpretation of assimilation using only ground in-situ measurements or ground and aircraft in-situ measurements (`G_V` and `GA_V`). While assimilating in-situ measurements from ground and aircraft together (`GA`) has the best match to the validation measurements, the other runs (`TneGA`, `TaGA`, `G`, `Ta`, `TaG`) have lower mean and median mismatch for all other comparisons, against the aircraft measurements (`A`), the ground-based in-situ measurements (`G`), all TCCON sites (`Ta`) and against the TCCON sites in Europe (`Te`), with rising difference in that order. However the comparison of runs assimilating other data against aircraft measurements shows a good match. This indicates that the mismatch when assimilating aircraft data does not stem from measurement errors in the aircraft measurements, but rather from the effect of their temporal and spatial distribution on the optimized field. The effect of assimilating aircraft measurements on the optimized fluxes is investigated in more detail in section 4.4.2.[5]Also, the very good match shown against the validation measurements when assimilating in-situ measurements from ground and aircraft needs to be seen as a warning about the effects of optimizing parameters for a predefined set of validation sites. This effect could be shown more prominently with resampling by site (the uncertainty of the median in Figure 4.8 only uses resampling by measurement).}

Assimilating non-European TCCON sites in addition to in-situ measurements from ground and aircraft (`TneGA`) reduces the mismatch to all measurements except for the validation sites. This points towards effects caused by information in higher layers of the atmosphere which skew the assimilation when using aircraft measurements, but get corrected by the information from TCCON measurements.

4.4.2 The effect of assimilating aircraft measurements when using a biased prior

The cross-comparison in section 4.4.1 show significant differences between assimilating in-situ aircraft data alongside in-situ ground data (`GA`) and assimilating only in-situ ground data (`G`), with a stronger mismatch when using aircraft data and validating against TCCON data (`GA_Ta`). These differences are most pronounced in Europe (`GA_Te`). A detailed look at the European TCCON site in Bremen, Germany, which is assimilated in none of the runs supports this result as shown in Figure 4.9. The figure shows histograms of measured total columns of CO_2 (tccon-bremen) and corresponding total column values retrieved from the modelled CO_2 field at the same time (Carbontracker). It includes a linear least squares fit to the data with a value dependence and an absolute shift. The run with a priori fluxes at the top left overestimates the total column measurements at the Bremen site with an average bias of almost -4 ppm, but with little dependence on the absolute value (3%). Using only the in-situ ground data (`G`) as shown in the middle left histogram improves the bias to about -0.38 ppm at 390 ppm, but with a dependence on the absolute value of -7%. Adding aircraft measurements in the top right histogram (`GA`) increases the value dependence to -9% and the bias at 390 ppm to -0.88 ppm. This shows that instead of improving the match to total column measurements by providing a constraint in high altitudes, aircraft measurements reduce this match in this study.

However, adding measurements from non-European TCCON sites (`TneGA`), as shown in the middle right, reduces the value dependence to -3% and the bias at 390 ppm to -0.64 ppm. The fit between

[5]{

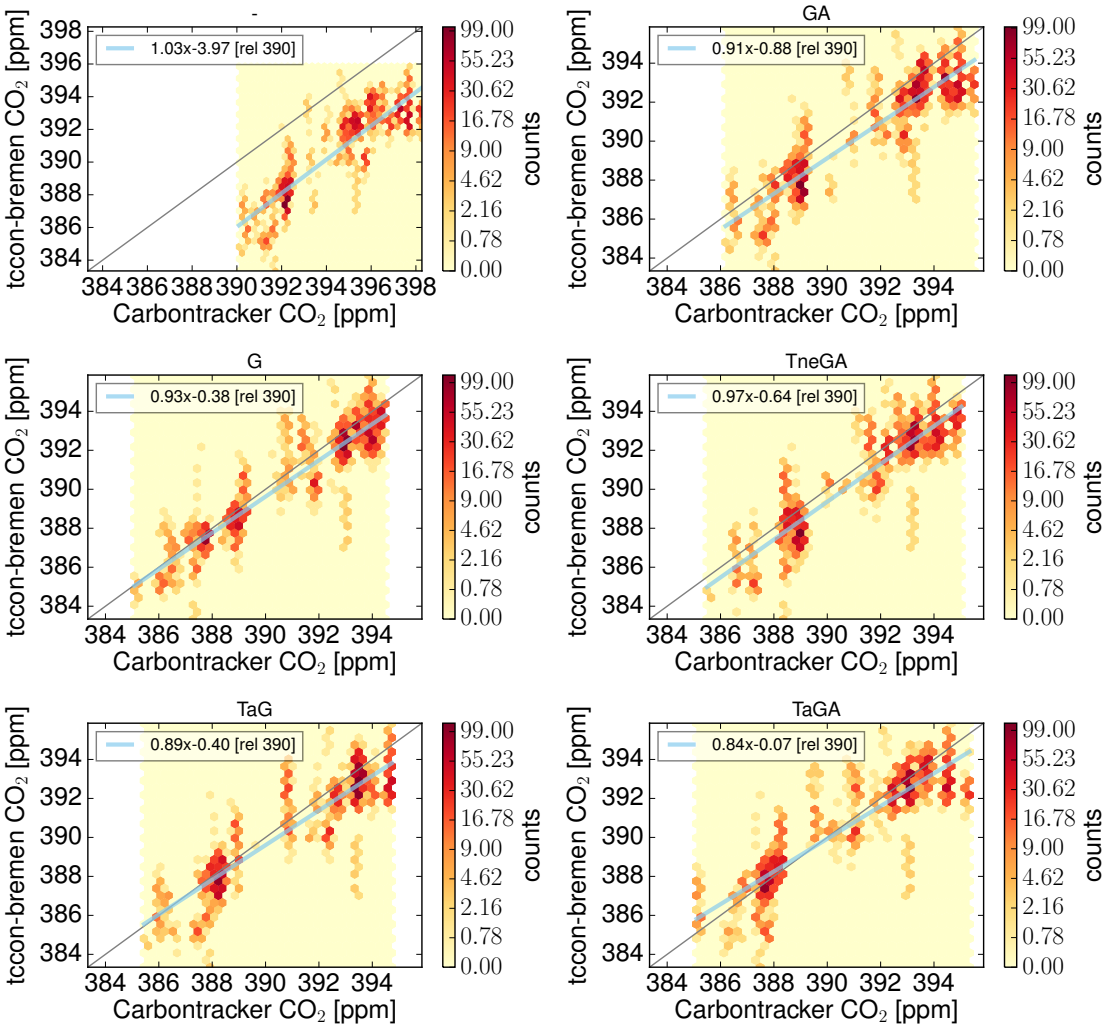
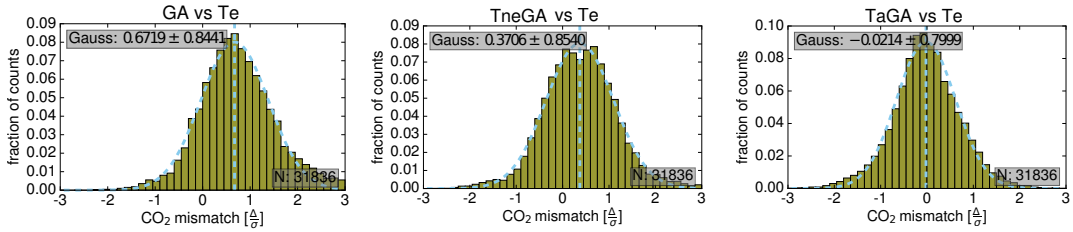


Figure 4.9: Scatter-histogram of CO_2 total column measurements from the TCCON site in Bremen vs. modelled total column data from Carbontracker. The bias in the fit is given as TCCON measurement minus model where the model estimates 390 ppm. The titles use the shorthand defined in table 4.1 to identify assimilated measurements.



(a) no TCCON (b) non-EU TCCON (c) all TCCON

Figure 4.10: Residuals between model and the TCCON sites in Europe when assimilating in-situ ground and in-situ aircraft data along with no TCCON data, non-European TCCON data or all TCCON data, normalized by inverse uncertainty and by site. The run in histogram (c) assimilates the validation sites, so it only verifies that when assimilating the sites, the modelled concentrations match the measurements at European TCCON sites. All histograms are calculated on a timeframe from 2010-12-03 to 2011-12-02.

the model and TCCON measurements is closer to the fit when using only the ground data (`G`), with a larger bias, but a lower dependence on the absolute value than when using aircraft data without TCCON (`GA`). This supports the finding from the cross-comparison in section 4.4.1 which used all European TCCON data, but did not discriminate between different sites. The histograms in Figure 4.10 show that this result holds in general, with the normalized (and therefore unitless) bias against European TCCON sites of 0.67 when using only in-situ data (`GA`) reduces to 0.37 when adding measurements from TCCON sites outside Europe (`TneGA`).[6] As such, assimilating non-European TCCON sites corrects flux attribution errors caused by assimilation of the aircraft measurements.

The plots at the bottom of Figure 4.9 (`TaG` and `TaGA`) show that when assimilating TCCON sites, the match to European TCCON measurements does not degrade by adding aircraft measurements. This is an indicator that the additional information from assimilating TCCON measurements helps avoiding mistaken adjustments due to assimilating the temporally sparse aircraft data.

As expected from the changed match to non-assimilated measurements, assimilating aircraft data has a noticeable effect on the fluxes. However, the changes are smaller than the two sigma significance level. Figure 4.11 shows the regional attribution of fluxes on continental scale for the different runs. Changes which are stronger than the one sigma uncertainty estimate are an increased sink in North America when adding TCCON data to the assimilation paired with a decreased sink in Atlantic and Pacific. For Eurasia, assimilation of TCCON data shows opposing effects whether aircraft in-situ measurements are assimilated (`TaGA`) or not (`TaG`). The runs assimilating only in-situ ground measurements (`G`) and assimilating in-situ ground and aircraft measurements along with TCCON measurements (`TaGA`), show roughly the same flux, while the run assimilating only in-situ ground and TCCON measurements (`TaG`) shows a more than 1 PgC/a weaker sink than the run using in-situ ground measurements without TCCON (`G`). These two runs also show no significant sink in South America, which casts doubt on their results. Therefore the first (`TaG`) is not used in section 4.5. The second (`G`) is used, because it follows the baseline run from section 3 and as such provides a comparison between the two different

[6] The histograms are normalized by site, ensuring that each European TCCON site has the same weight in the comparison, despite different measurement frequencies.

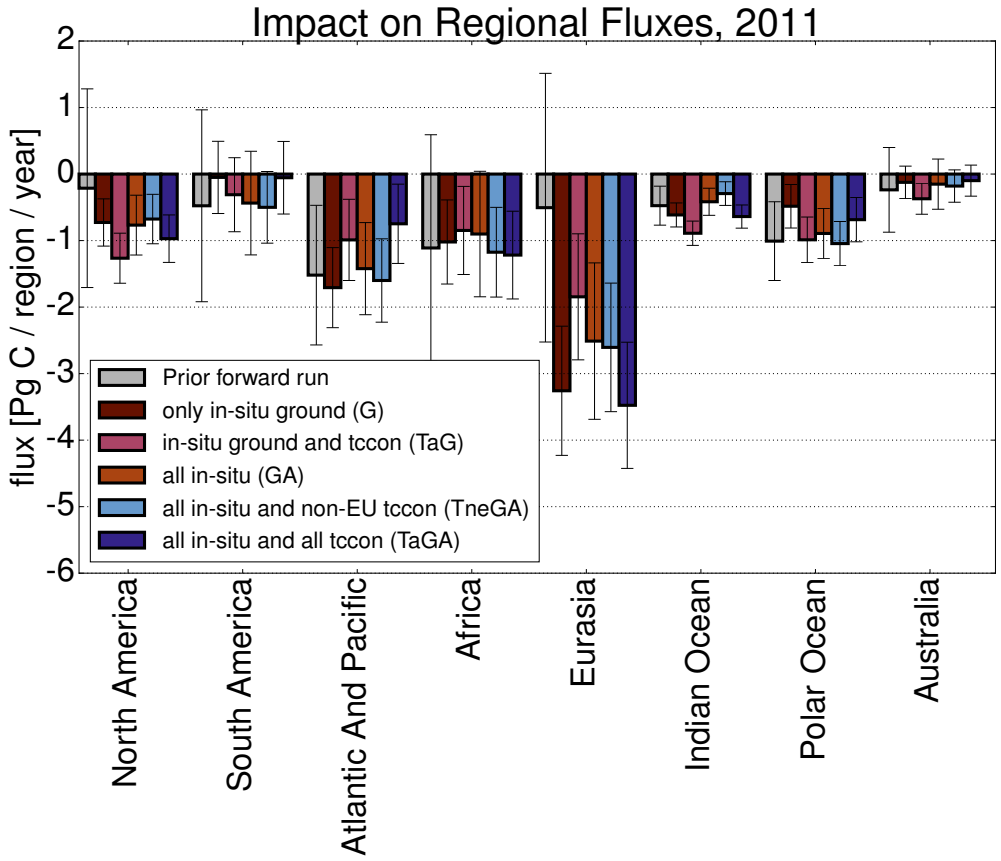


Figure 4.11: Aggregated fluxes from 2010-12-03 to 2011-12-02 for the different runs shown in Figure 4.9.

CarbonTracker versions used: gridded here, non-gridded in section 3. Both versions are described in section 2.5.3.

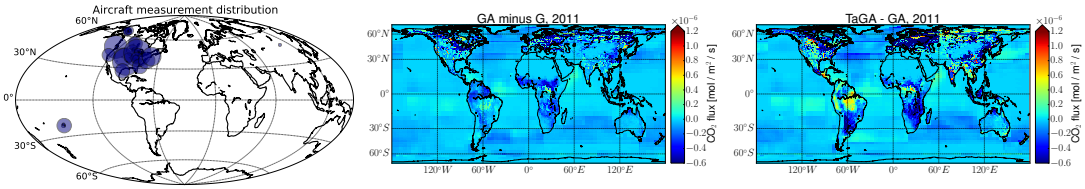
Assimilating non-European TCCON sites in addition to aircrafts does not show significant effects on continental fluxes, despite the effect on validation at European TCCON sites shown in Figure 4.10. There are changes in all the different regions but they are clearly below the one-sigma significance criterion.

In short, the best estimates of non-assimilated observations are achieved by assimilating in-situ ground, in-situ aircraft and TCCON measurements together (`TaGA`).

However, the reduced sink in Atlantic and Pacific seen in Figure 4.11 when assimilating TCCON data hints at a relationship between flux attribution and the location of measurement sites. This is investigated in section 4.4.3.

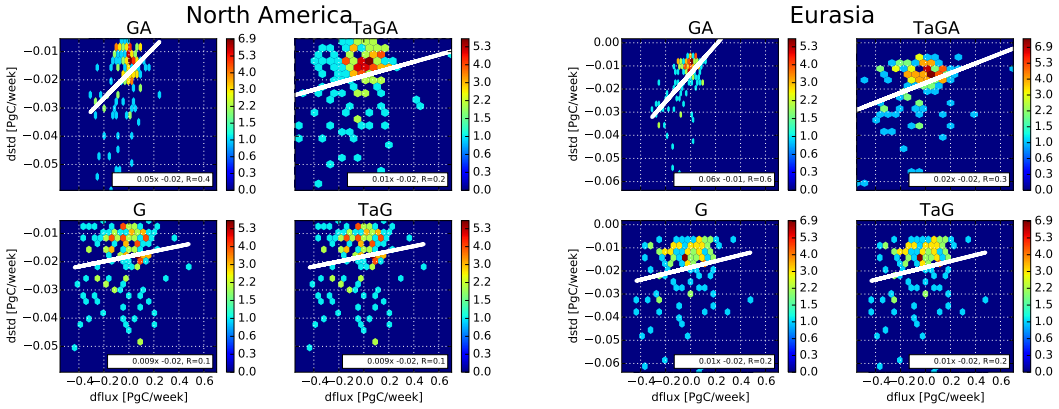
4.4.3 Unequal observational pressure guides flux attribution

When measurements are assimilated, the uncertainty of adjusted fluxes is reduced for purely mathematical reasons. As such, the reduction in uncertainty in a specific region can be used as a



(a) aircraft meas dist (b) adding aircrafts, GA - G (c) adding TCCON, TaGA - GA

Figure 4.12: Distribution and density of assimilated aircraft measurements.[7] The flux maps show differences between assimilating both aircraft and ground in-situ measurements `GA` and only in-situ ground measurements (`G`, 4.12b) and between assimilating aircraft, ground and TCCON (`TaGA`) and both in-situ ground and aircraft (`GA`, 4.12c)..



(a) North America (b) Eurasia

Figure 4.13: Correlation histograms between flux change (dflux) and uncertainty reduction (dstd) in Eurasia and North America (aggregated). These show the correlation between changes in the flux and reduction in uncertainty compared to the prior, including fit parameters of a linear fit and its correlation coefficient `R`. The titles use the shorthand defined in table 4.1 to identify assimilated measurements.

proxy for pressure from observations: the effect from assimilating measurements on the modelled fluxes.

As shown in Figure 4.12a, most assimilated aircraft measurements are located in North America. Figures 4.12b and 4.12c indicate that where adding aircrafts (`GA - G`) adds a source, adding TCCON to this (`TaGA - GA`) adds a sink, the inverse change. Adding TCCON measurements (`Ta`) generally reverses the global shifting of fluxes seen when adding aircraft measurements (`A`). This effect is most visible in South America, Africa and Eurasia.

An explanation of these shifts in global flux distribution is that measurements from aircrafts give a higher weight to flux adaptation in North America, since the fluxes there are closest to the aircraft measurements. Due to the (known) underestimation of the global sink in the prior fluxes, this adaptation could result in an increased sink in North America. Adding TCCON measurements in Europe shifts the region of highest measurement pressure partially back towards Eurasia, resulting in a shift of the global sink. Such an effect would cast in doubt flux attribution from models

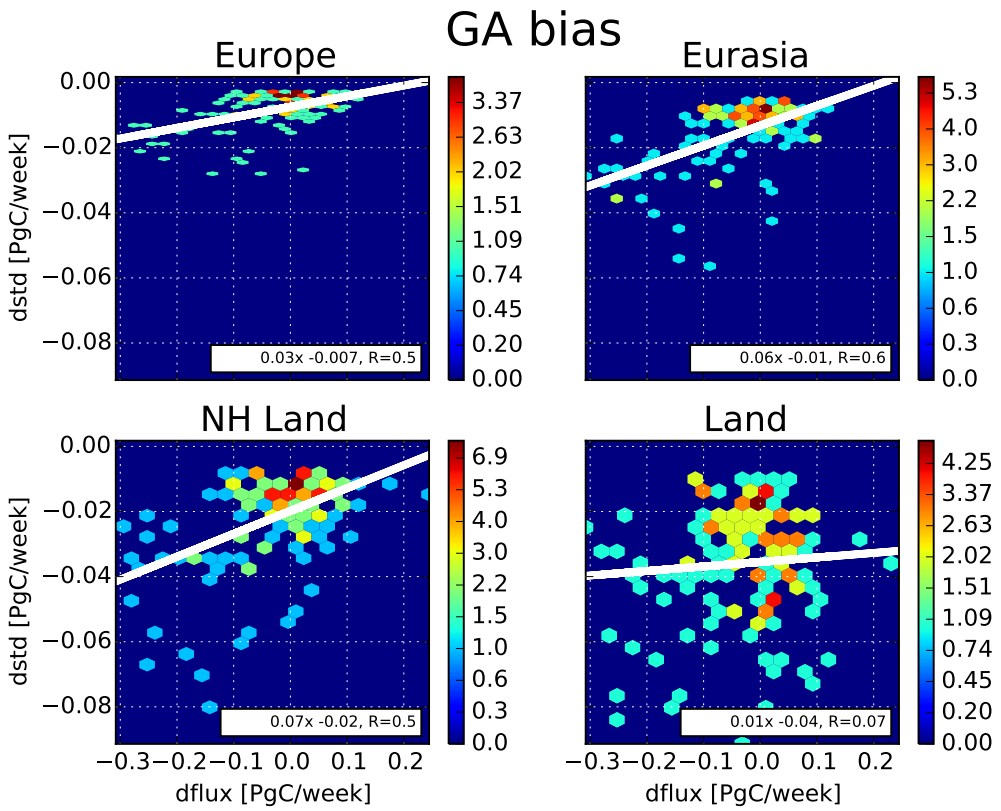
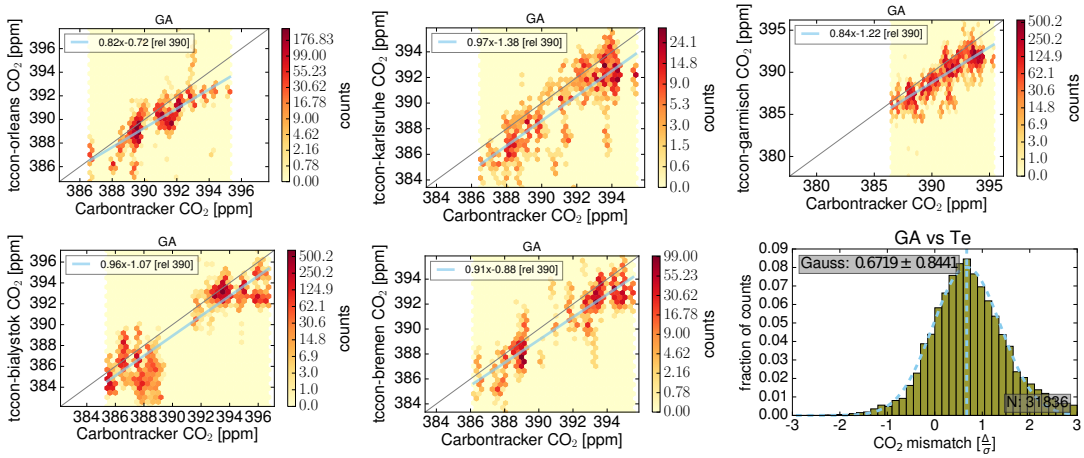


Figure 4.14: Correlation histograms between flux change (dflux) and uncertainty reduction (dstd) for regions of different scales (aggregated) when assimilating in-situ measurements from ground sites and aircraft (`GA`). These show the correlation between changes in the flux and reduction in uncertainty compared to the prior, including fit parameters of a linear fit and its correlation coefficient `R`. NH Land contains transcom regions for North America, Asian Temperate, Asian Boreal, Asian Temperate and Europe, but excludes tropical land regions (Tropical Africa and tropical Asia).

with biased prior. It would show up in a correlation between flux uncertainty reduction and an increased sink.

Figure 4.13 shows the flux change in North America and Eurasia correlated with the change of the uncertainty in the region and time compared to the prior. The plot contrasts the different runs in-situ ground (`G`), in-situ ground and aircraft (`GA`) and both these runs with the addition of TCCON data (`TaG` and `TaGA`). It shows a correlation between reduction of uncertainty and flux in Eurasia and North America when assimilating aircraft data: On average reduced uncertainty moves the flux towards a sink, because the prior has a too weak sink and it is corrected where there are measurements. Adding assimilation of TCCON measurements increases the spread of the flux adaptation, reducing the correlation compared to a linear fit, but the points in the histogram which give rise to the correlation when only assimilating in-situ data stay visible by eye.



(a) Orleans (top) Bialystok (bottom)

(b) Karlsruhe (top) Bremen (bottom)

(c) Garmisch (top) mismatches (bottom)

Figure 4.15: Scatter-histograms of CO_2 total column measurements from European TCCON sites vs. modelled total column data from Carbontracker using a run assimilating in-situ measurements from ground sites and aircrafts. The bias in the fit is given as TCCON measurement minus model where the model estimates 390 ppm. The histogram shows the mismatch against the five sites normalized by site and inverse uncertainty, calculated from 2010-12-03 to 2011-12-02.

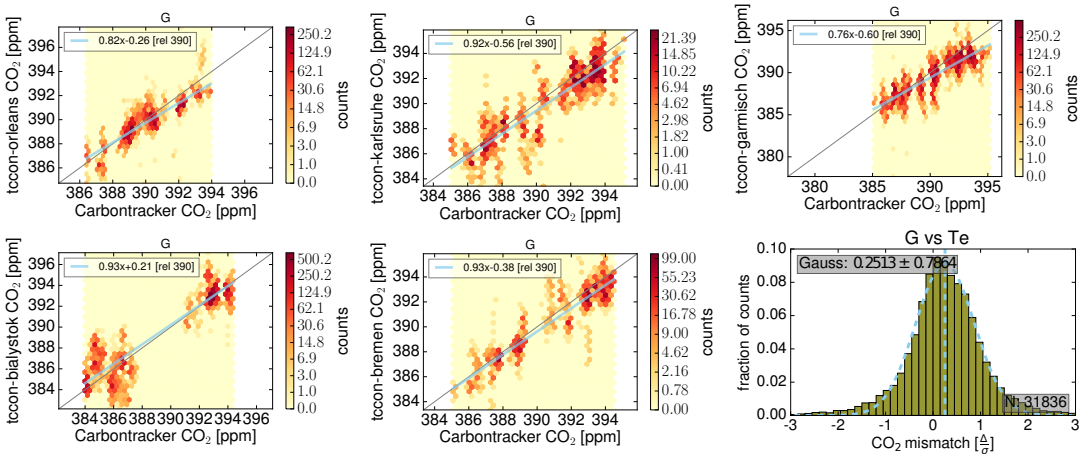
As shown in Figure 4.14, increasing the size of the region to all land on the northern hemisphere preserves the correlation, but extending it to also include the land on the southern hemisphere makes it disappear. Unequal observational pressure even affects the carbon balance between the hemispheres, but much less the global balance between land and oceans.

This result highlights the importance of preconditioning the prior fluxes. Further it allows defining a criterion for good preconditioning of the prior fluxes: there should be no correlation between the change in flux uncertainty and the change in the flux. To find a robust flux distribution on regional scale, the flux prior should be adjusted to show no global bias. Even though a global bias of the fluxes is corrected by a small number of observations, as shown in the observation density experiment in section 3.2.2.2 and the assimilation of TCCON measurements with high uncertainty in section 4.3, Figure 4.4, this can lead to regional and temporal misattribution of fluxes in the inverse direction of the global bias of the prior.

Since the global land flux correction shows no strong correlation between uncertainty reduction and the direction of the flux correction (Fig 4.14), using a globally aggregated weekly flux correction from an initial run of Carbontracker to precondition the prior fluxes should introduce no additional misattribution.

4.5 A new limit for the uncertainties of Eurasian fluxes

Using Satellite measurements, Reuter et al. (2014) estimated a European sink of 1.02±0.30 GtC/a for 2010, where Babenhauserheide et al. (2015) found a sink of 0.66±0.70 PgC/a with Carbontracker and 0.52±0.32 PgC/a with TM5-4DVar for April 2009 to April 2010. Knowing

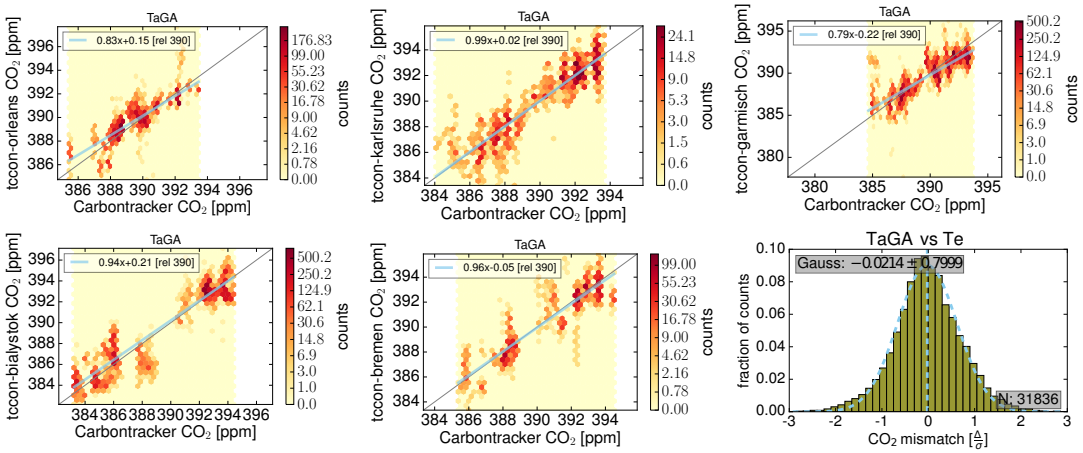


(a) Orleans (top) Bialystok (bottom)

(b) Karlsruhe (top) Bremen (bottom)

(c) Garmisch (top) mismatches (bottom)

Figure 4.16: Scatter-histograms of CO_2 total column measurements from European TCCON sites vs. modelled total column data from Carbontracker using a run assimilating in-situ measurements from ground sites but no aircrafts. The bias in the fit is given as TCCON measurement minus model where the model estimates 390 ppm. The histogram shows the mismatch against the five sites normalized by site and inverse uncertainty, calculated from 2010-12-03 to 2011-12-02.

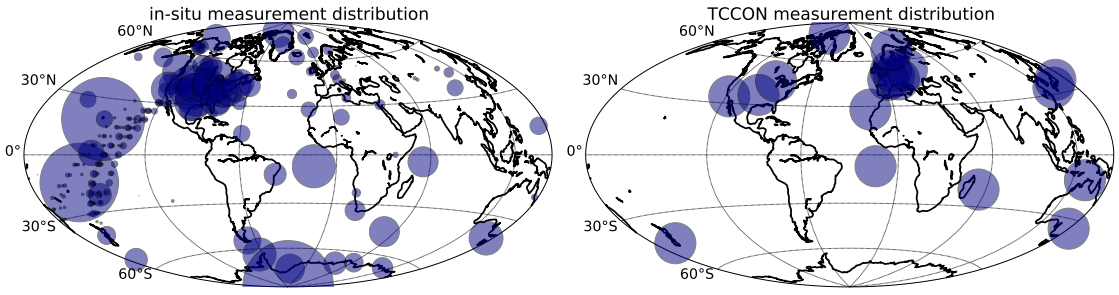


(a) Orleans (top) Bialystok (bottom)

(b) Karlsruhe (top) Bremen (bottom)

(c) Garmisch (top) mismatches (bottom)

Figure 4.17: Scatter-histograms of CO_2 total column measurements from European TCCON sites vs. modelled total column data from Carbontracker using a run assimilating the most recent TCCON data which contains the Bremen site and includes the fixes by Matthias/Hase. Those fixes resulted in differences up to 1 ppm in Karlsruhe, Germany by correctly fitting the background of the spectrum. The bias in the fit is given as TCCON measurement minus model where the model estimates 390 ppm. The histogram shows the mismatch against the five sites normalized by site and inverse uncertainty.



(a) in-situ (b) TCCON

Figure 4.18: Measurement density from in-situ sites (ground and aircraft) and TCCON sites.

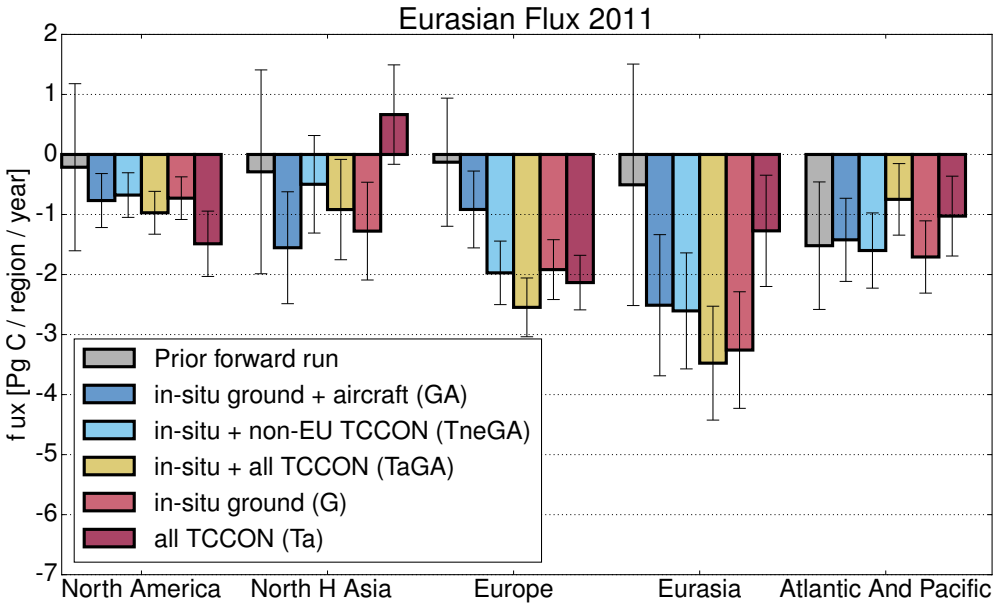


Figure 4.19: Fluxes aggregated from 2010-12-03 to 2011-12-02 for the given regions, with North H Asia combining Eurasian Boreal and Temperate and Eurasia combining North H Asia, Europe and Tropical Asia.

that the TCCON measurements correct problems for European concentration fields caused by assimilating aircraft measurements allows revisiting this question.

However, there are few in-situ measurements in Europe which together with total column measurements would allow discriminating remote fluxes from local fluxes. Even with the European TCCON sites, the assimilation still lacks measurements in central and boreal Asia to distinguish between Asian and European fluxes. Therefore the focus will be the whole Eurasian continent, because the European sink and the Asian sink cannot clearly be distinguished with this measurement network.

Figure 4.19 shows continental scale flux estimates from the runs with different assimilated observations: in situ ground and aircraft (`GA`), in-situ ground and aircraft with non-European

TCCON sites (`TneGA`), in-situ ground and aircraft with all TCCON sites (`TaGA`), in-situ ground (`G`) and all TCCON sites without any in-situ measurements (`Ta`).

Several caveats apply to these flux estimates. As shown in section 4.4.2, while the fluxes in Eurasia are strongly affected by assimilating European TCCON measurements, attribution of fluxes is partly guided by the measurement density. The run assimilating only TCCON measurements (`Ta`) shows strong deviations from the flux in North America, which was shown in section 3 as robust region due to the high density of observation sites. Other fluxes from `Ta` are likely contaminated by fluxes compensating for the stronger North American sink. Finally, the run assimilating only in-situ measurements from ground and aircraft (`GA`) is contaminated by strong guidance of flux attribution by measurement density which are corrected when assimilating TCCON sites outside Europe (shown in section 4.4.2 and Figures 4.15 and 4.17).

Therefore the flux results to use for investigating the uncertainty of flux estimates due to characteristics of the measurements stem from three runs: only assimilating in-situ ground measurements (`G`), assimilating in-situ ground, aircraft and non-European TCCON measurements (`TneGA`) and assimilating all available measurements (`TaGA`). As shown in Figure 4.18, these runs have an even distribution of measurement sites: The high density of aircraft measurements in North America (see Figure 4.12a) is either avoided or compensated by additional sites in other regions.

This allows estimating the uncertainty of the fluxes due to the characteristics of assimilated measurements. The Eurasian sink differs between 2.5±1 PgC/a for assimilating in-situ and non-European TCCON sites (`TneGA`) and 3.5±1 PgC/a for assimilating in-situ and all TCCON sites (`TaGA`) in-situ ground assimilation (`G`) in between. The European fluxes of these runs are spread between 2 PgC/a and 2.5 PgC/a with a one sigma uncertainty of ±0.5 PgC/a. The estimate is within two sigma of the fluxes reported by Reuter et al. (2014) who assimilated satellite retrievals.

Adding non-European TCCON sites (`TneGA`) does not significantly change the flux in Eurasia, but rather redistributes the sink from temperate Asia towards Europe (see Fig. 4.19). Adding European TCCON sites (`TaGA`), however, increases the aggregated sink in Eurasia, while decreasing the sink in Atlantic and Pacific. The sink in Europe is similar whether assimilating in-situ sites with non-European TCCON sites (`TneGA`), in-situ sites and all TCCON sites (`TaGA`) or only TCCON sites (`Ta`). However, assimilating only in-situ ground measurements (`G`, 4.16) gives a similarly good match to European TCCON sites as assimilating all measurements (`TaGA`, Fig. 4.17), while assimilating only TCCON measurements (`Ta`) still shows a good match against ground sites (Fig 4.4). This points towards a stronger sink in Europe than seen by either Reuter et al. (2014) or Babenhauserheide et al. (2015), but highlights again that proper flux attribution within Eurasia requires higher density of measurements. Another clear indication that the density of measurements from using only TCCON measurements (`Ta`) does not suffice to resolve sub-continental fluxes in Eurasia is that north hemispheric Asia is estimated to be a source, though it includes the boreal forests.[8] However, the total Eurasian flux is still within two sigma of the 2.5 PgC flux shown in Figure 3.10.

In general assimilating in-situ ground measurements and TCCON measurements together reduces the uncertainty of the fluxes while keeping a similarly good fit to aircraft measurements, and assimilating all measurements gives a best fit to the Bremen TCCON site (not assimilated in the

[8] The attribution within Eurasia when using only TCCON measurements is less robust than when also assimilating in-situ measurements, because as shown in Figure 4.18, the in-situ measurements include four sites in eastern Asia which are missing in the pure TCCON run (validation of the TCCON-only run against these sites does not show major problems, though).

comparison run in Figure 4.9), suggesting that the fluxes provide the best match to atmospheric concentration measurements. Still, the fluxes for which the model assimilated as much information as possible suffer from insufficient (essentially in-existent) independent validation.

As shown in 4.19 the plausible range of Eurasian fluxes spans from 0.5 Pg carbon per year up to 4 Pg carbon per year. It is important to note that the assimilation in this chapter uses the gridded version of Carbontracker as described in section 2.5.3. Consequently all the results are preliminary.

4.6 Summary

This chapter presents a joint assimilation of in-situ and total column measurements to estimate the global distribution of CO_2 surface fluxes. Measurements with different spatial and temporal structure yield orthogonal information which improves the robustness of the estimated sources and sinks.

Based on more than 3 years of assimilated measurements, this part of the study assesses the impact of assimilating different kinds of measurements into the CarbonTracker data assimilation system. The system assimilates total column data from the TCCON network and in-situ measurements from the obspack compilation, the latter from ground and from aircraft.

It is shown that (a) assimilating high precision total column data from 18 TCCON sites achieves a comparable level of quality as assimilating data from over 100 in-situ sites from the obspack compilation, that (b) assimilating in-situ data from aircraft measurements skews the CO_2 concentration field in Europe, leading to a higher mismatch against European TCCON sites which is reduced when additionally assimilating non-European TCCON sites and that (c) the location of measurement sites skews the attribution of the global sink towards regions with high observation pressure. Avoiding this effect will either need adjustment of representativeness errors per region to equalize globally the observation pressure, or require a correction of global biases in the prior fluxes before using the prior fluxes in the assimilation.

Investigating the resulting fluxes from assimilating different sets of observations points towards a stronger European sink than previous estimates which only used in-situ measurements. Feng et al. (2015) raise the question whether the larger European sink found by Reuter et al. (2014) when using satellite measurements is a real signal or an artefact of the inversion. The results shown in Figure 4.19 as well as the reduced misattribution due to correlation of observation pressure and flux adaption (section 4.4.3) support the conclusion that the stronger biospheric sink in Europe seen when assimilating total column measurements is a real signal. The uncertainty caused by characteristics of the measurements is on the order of the model-internal uncertainty estimate.

Chapter 5

Conclusions and Outlook

Inverse modelling of carbon dioxide surface fluxes has become an important approach to gain insight into the functioning of the Earth's contemporary carbon cycle. Particular interest relates to the land biosphere and the oceans which, on global scale, take up roughly half of the carbon dioxide emitted by human activity. On smaller scales such as continents or countries, however, it is highly uncertain whether the large gross fluxes across the land-atmosphere and ocean-atmosphere interface balance to a net sink or source of carbon dioxide within a given region. Uncertainties become even larger when projecting the future evolution of the natural carbon cycle under climate change.

This study aims at contributing to a better understanding of the capabilities of current CO_2 inverse modelling, what uncertainties to expect and how to advance toward robust regional-scale surface flux estimates. To this end, the study first compares two established inverse models for CO_2 surface flux estimation gaining insight into modelling uncertainties and sensitivities. Second, the study examines the benefit of simultaneously assimilating different types of measurements.

In the first part – the comparison of the inverse models – both models are run in a harmonized setup (Babenhauserheide et al., 2015). Assimilated measurements and prior fluxes are chosen identical. Both inverse models use the same transport model and are evaluated consistently. Finally the background covariance is harmonized as far as the spatially and temporally different flux representations permit.

The selected inverse models are CarbonTracker and TM5-4DVar. Both models approximate the optimal solution to make the inversion viable. The different approximations lead to different deviations from the ideal solution. Thus inter-model differences can serve as estimate of the uncertainty due to the choice of approximations.

CarbonTracker uses an ensemble method which steps through the assimilation time period sequentially. It scales fluxes binned spatially in regions with similar ecosystem (ecoregions) and binned temporally in a time window of 5 weeks, progressing by one week in each step. TM5-4DVar uses a four dimensional variational method and adjusts the flux following the largest singular values one by one over the whole assimilation time period. TM5-4DVar uses spatial binning on a $6^\circ \times 4^\circ$ longitude×latitude grid and temporal binning in monthly fluxes.

The assimilation experiments performed with both models provide a measure of inter-model uncertainties due to from differences in the inverse method und the flux representation. Comparing the optimized atmospheric CO_2 concentration fields from both models to non-assimilated observations shows that the difference between the models is lower than the uncertainty due to

the limited number of non-assimilated observation sites. Therefore the quality of the estimated fluxes is indistinguishable and it is possible to use the difference between the fluxes as a lower limit for the flux uncertainty. Additionally the criteria for selecting a method reduce to practical considerations, such as the complexity of adapting the method to other transport models.

Varying the number of assimilated observations based on historical availability reveals a dependency between the lower limit for the uncertainty from inter-model differences and the observation coverage. For continental-scale biospheric net fluxes the flux uncertainty ranges from 0.1 to 0.5 petagram carbon per year when assimilating all currently available in-situ measurements. This is about 25% to 100% of the continental biospheric sink. Adding more measurements decreases the lower limit of the uncertainty, with North America, where the density of in-situ measurements is highest, reaching the lowest uncertainty of 0.1 petagram carbon per year. This shows that the main limitation for the accuracy of the estimated surface fluxes due to inter-model differences is the density of measurement sites.

In addition to the effects of limited measurement density, comparison of flux estimates in Asia reveals uncertainties of the fluxes due to the temporal binning. In Asia, the differences between the prior fluxes and the fluxes optimized by CarbonTracker vary significantly week over week. These corrections cannot be represented with monthly binned adaptations to fluxes as used by TM5-4DVar. This leads to flux artefacts in TM5-4DVar which accumulate to 0.5 Pg carbon per year. The artefacts are reproduced in CarbonTracker when coarsening its temporal binning from 7 days to 20 days, but are not detectable with the regular weekly binning.

To match summer concentrations of CO_2 in Antarctica with CarbonTracker, its assimilation time window must be increased from 5 weeks to 10 weeks: the assimilation time window must be long enough to connect regions with high biospheric activity to measurement sites. Carbon transported from these regions needs to reach atmospheric measurement sites by means of atmospheric transport within that time period.

In the second part of the study, total column measurements of CO_2 from the ground-based TCCON network are assimilated in CarbonTracker. The TCCON network was selected because it provides the most accurate total column measurements. CarbonTracker was chosen for the practical reason that it can be used with transport models which do not provide an adjoint, therefore it will be easier to adapt to the transport models used and developed at the Institute of Meteorology and Climate Research at KIT. This second part uses a newer version of CarbonTracker which optimizes fluxes on a $1^\circ \times 1^\circ$ longitude×latitude grid and recreates the ecoregion structure of the previous version by adding strong correlations between grid-boxes in the same ecoregion.

Assimilating measurements with different sensitivities und coverage allows investigating the impact of the characteristics of different kinds of measurements on the estimated fluxes. The investigation yields two key findings: First, assimilating in-situ measurements from aircrafts together with ground-based in-situ measurements leads to misattribution of the fluxes on continental scales. This misattribution is corrected when adding total column measurements. However, while it is strongest with aircraft measurements, such a misattribution cannot be completely avoided when the flux prior has a global bias. Second, the best estimate of Eurasian fluxes is achieved by assimilating in-situ measurements from ground and from aircraft together with total column measurements. This yields a one-sigma range of the Eurasian sink between 2.5 Pg and 4.5 Pg carbon per year, using the model estimate of the uncertainty. The additional uncertainty for Eurasian fluxes due to the temporal and vertical coverage of measurements is on the order of 1 Pg carbon per year.

The assimilated aircraft measurements are only available with irregular temporal distribution. For most locations, there are measurements for only a few days per year. Assimilating these measurements leads to adjustments of the CO_2 concentrations in the middle and upper troposphere, revealed by mismatches to total column measurements. These concentrations in higher regions of the atmosphere only have a small impact on ground based in-situ measurements, though, which causes them not to be corrected by in-situ measurements conducted on the ground. Consequently, constraints from the aircraft measurements which are only representative for a short time cause lasting concentration changes and misattribution of fluxes.

A further effect contributing to the observed mismatch between modelled concentration fields and total column measurements is unequal observational pressure. Fluxes are adjusted to fit the measurements. If the measurement density is highest in one given region, the flux adjustment reduces mismatches to the measurements in this region more strongly than for other regions. If the prior fluxes have a global bias, the correction for this bias – in case of this study an increase of the biospheric carbon sink – is localized in the region with the highest density of measurements. With aircraft measurements this region currently is North America. In short: The global biospheric sink is attributed to the region with most measurement sites. Assimilating total column measurements in addition to in-situ measurements equalizes the density of measurement sites and as such the observation pressure, since the density of total column sites is highest in Europe.

The irregular temporal distribution of aircraft measurements and the global bias of the prior together lead to an overestimation of European CO_2 total column concentration measurements by about 1 ppm when assimilating aircraft measurements in addition to ground based in-situ measurements. Assimilating total column measurements outside Europe reduces this concentration bias to about 0.5 ppm. Also adding total column measurements within Europe reduces the concentration bias of the non-assimilated TCCON site in Bremen to 0.07 ppm. This shows that assimilating total column measurements corrects the artefact introduced by the assimilation of aircraft measurements.

For the Eurasian sink, the one sigma range between 2.5 Pg and 4.5 Pg carbon per year is larger than reported in previous studies – including the results from Reuter et al. (2014) who used total column measurements from satellite – and larger than those found in the comparison from the first part of this study which used only in-situ data. All fluxes are still within their reported two sigma uncertainty, though. To avoid artefacts from unequal observation pressure, only the assimilation which uses in-situ data from ground and the assimilation which uses all TCCON measurements in addition to in-situ measurements from ground and from aircraft are used to estimate the Eurasian fluxes.

Taken together, the comparison of two models in the first part and the comparison of assimilation using different kinds of measurements in the second part provide an estimate of the magnitude and uncertainty of the biogenic sink in Europe. The differences between results can be used to calculate an aggregated estimate for the systematic error. The inter-model difference found is below 0.2 Pg carbon per year and the difference from characteristics of the measurements is on the order of 0.5 Pg carbon per year. These differences are consistent with an unbiased standard deviation of 0.4 Pg and 0.9 Pg, respectively. The flux is taken from the setup which assimilates all measurements. Adding the model-intrinsic uncertainty of 0.5 Pg carbon per year, this gives a best estimate of the European biogenic sink of 2.5 ± 1.8 Pg carbon in 2011.

Feng et al. (2015) raised the question whether this stronger European sink seen when assimilating total column measurements is a real signal or an artefact. Independent reproduction in this study of the stronger sink seen by Reuter et al. (2014) indicates that this sink is likely a real signal.

However comparing CO_2 concentrations from different fluxes with the measurements in central Asia did not find a robust difference between the different setups. As such the question is still open, due to missing measurements with sufficiently high frequency and accuracy in central Asia.

Outlook: This study showed that improving the robustness of the estimated sources and sinks of carbon mostly requires assimilating additional measurements, especially in weakly constrained regions like the southern hemisphere. It showed that assimilating total column measurements can correct artefacts introduced by the temporally sparse aircraft measurements, complementing in-situ measurements. This indicates that temporally and spatially sparse satellite measurements should likewise complement the currently assimilated measurements. However the correction of these artefacts works only partially for regions for which no total column measurements are available. Therefore a future strategy to find more robust flux estimates should improve the situation using all four of the following angles: increasing the global coverage of in-situ sites, especially outside the USA, increasing the frequency of aircraft measurements to reduce the artefacts incurred by their temporal structure, improving the global coverage of total column measurement sites to correct the artefacts from assimilating aircraft measurements, and assimilating satellite measurements to improve the coverage in inaccessible regions.

Also integrating different meteorological models and different kinds of measurements presents a challenging scientific goal, since this integration requires careful investigation of the characteristics of the different methods and measurements.

To avoid misattribution of the fluxes due to unequal observation pressure of the available measurements, the prior fluxes should be corrected to remove a global bias of the net flux before using the prior fluxes in the assimilation.

As a concrete step, a robust estimate of the biospheric carbon sink in Europe would require at least one additional measurement site in central Asia. Having an elevated in-situ instrument (i.e. a tower site) and a remote sensing instrument at the same location would allow distinguishing between local biosphere and remote fluxes, which together with European sites should provide the means to calculate robust European fluxes.

An obvious path towards extension of the total column measurement network would be to build on the work of Barthlott et al. (2015) and investigate whether the CO_2 total column measurements conducted by many sites in the Network for Detection of Atmospheric Composition Change (NDACC) provide sufficient accuracy to use as constraint for estimating surface fluxes. The measurements from NDACC go back 20 years, so they would improve long-term statistics, making it easier to detect changes in the sources and sinks of carbon dioxide. Improved flux estimates can inform decision making, because the carbon sink in the biosphere and the oceans determines the fraction of antropogenic carbon emissions which stays in the atmosphere, the strongest anthropogenic driver of climate change.

The uncertainty this study finds for the net flux from different regions with the current measurement network makes quantification of biospheric carbon sources and sinks difficult, though. Providing estimates of the sources and sinks of carbon dioxide with sufficient accuracy to detect changes in the biosphere due to global warming would be even more challenging.

However, the key to a better scientific understanding of biospheric and oceanic flux estimates lies in assimilating more measurements at additional observation sites, especially in central Asia and on the southern hemisphere, as well as measurements from different types of instruments, like total column retrievals from ground, as shown here, or from satellite, for which the foundation was lain with the assimilation of ground-based total column measurements implemented in this study.

Chapter 6

Acknowledgements

Heartfelt thanks goes to my supervisor André Butz for his invaluable support throughout my PhD. Furthermore I thank Frank Hase for fruitful discussions on the viability of greenhouse gas observation from ground, Udo Grabowski for help whenever something did not work with our institutes cluster and for inspiring discussions about programming, Julian Kostinek for deep discussions which led me to a clearer understanding of technical and programming questions, Friedrich Klappenbach and Philipp Hahne for sharing insights, for help and for discussing whatever problem came up. Wouter Peters, Sourish Basu and Sander Houweling helped with inverse modelling. My Professors Peter Braesicke and Johannes Orphal provided encouragement and advice. Also I thank the other members of the RemoteC group, the collegues at IMK-ASF and especially the afternoon tea circle for their warm welcome, friendly collaboration and thought-provoking discussions.

André Butz and Arne Babenhauserheide are supported by the Emmy-Noether program of the Deutsche Forschungsgemeinschaft (DFG) through grant BU2599/1-1 (RemoteC). This study would not have been possible without the work of Maarten Krol (SRON Netherlands Institute for Space Research, Utrecht, the Netherlands; Institute for Marine and Atmospheric Research (IMAU), Utrecht University, Utrecht, the Netherlands; Department of Meteorology and Air Quality (MAQ), Wageningen University and Research Centre, Wageningen, the Netherlands) who maintains the Transport Model 5 (TM5). Also I am indebted to the many individuals and institutions who have contributed original data to the ObsPack project, specifically version v1.0.2 2013-01-28 used in this study.

TCCON data is available from the TCCON project. The work of all the individuals who provided data to the project makes assimilation of total column measurements feasible.

Finally I thank Paul Tol from SRON for his work on colourschemes which work for colour-blind people, O. Tange for easy ways to save time using GNU parallel (Tange, 2011) and E. Jones and J. D. Hunter for their tools for convenient data analysis with Python.

Chapter 7

Bibliography

Alexandrov, G. A.: Explaining the seasonal cycle of the globally averaged CO_2 with a carbon-cycle model, Earth System Dynamics, 5, 345–354, doi:10.5194/esd-5-345-2014, URL http://www.earth-syst-dynam.net/5/345/2014/, 2014.

Andrews, A. E., Kofler, J. D., Trudeau, M. E., Williams, J. C., Neff, D. H., Masarie, K. A., Chao, D. Y., Kitzis, D. R., Novelli, P. C., Zhao, C. L., Dlugokencky, E. J., Lang, P. M., Crotwell, M. J., Fischer, M. L., Parker, M. J., Lee, J. T., Baumann, D. D., Desai, A. R., Stanier, C. O., de Wekker, S. F. J., Wolfe, D. E., Munger, J. W., and Tans, P. P.: CO_2, CO and CH_4 measurements from the NOAA Earth System Research Laboratory's Tall Tower Greenhouse Gas Observing Network: instrumentation, uncertainty analysis and recommendations for future high-accuracy greenhouse gas monitoring efforts, Atmospheric Measurement Techniques Discussions, 6, 1461–1553, doi:10.5194/amtd-6-1461-2013, URL http://www.atmos-meas-tech-discuss.net/6/1461/2013/, 2013.

Arrhenius, S.: On the influence of carbonic acid in the air upon the temperature of the ground., vol. 5, pp. 237–276, The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, 1896.

Babenhauserheide, A., Basu, S., Houweling, S., Peters, W., and Butz, A.: Comparing the CarbonTracker and TM5-4DVar data assimilation systems for CO_2 surface flux inversions, Atmospheric Chemistry and Physics, 15, 9747–9763, doi:10.5194/acp-15-9747-2015, URL http://www.atmos-chem-phys.net/15/9747/2015/, 2015.

Baldocchi, D. D.: Global Change Biology, vol. 9, chap. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future, pp. 1–14, Blackwell Publishing Ltd, 2003.

Barthlott, S., Schneider, M., Hase, F., Wiegele, A., Christner, E., González, Y., Blumenstock, T., Dohe, S., García, O. E., Sepúlveda, E., Strong, K., Mendonca, J., Weaver, D., Palm, M., Deutscher, N. M., Warneke, T., Notholt, J., Lejeune, B., Mahieu, E., Jones, N., Griffith, D. W. T., Velazco, V. A., Smale, D., Robinson, J., Kivi, R., Heikkinen, P., and Raffalski, U.: Using XCO_2 retrievals for assessing the long-term consistency of NDACC/FTIR data sets, Atmospheric Measurement Techniques, 8, 1555–1573, doi:10.5194/amt-8-1555-2015, URL http://www.atmos-meas-tech.net/8/1555/2015/, 2015.

Basu, S., Houweling, S., Peters, W., Sweeney, C., Machida, T., Maksyutov, S., Patra, P. K., Saito, R., Chevallier, F., Niwa, Y., Matsueda, H., and Sawa, Y.: The seasonal cycle amplitude of

total column CO2: Factors behind the model-observation mismatch, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD016124, 2011.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO_2 fluxes estimated from GOSAT retrievals of total column CO_2, Atmospheric Chemistry and Physics, 13, 8695–8717, doi:10.5194/acp-13-8695-2013, URL http://www.atmos-chem-phys.net/13/8695/2013/, 2013.

Bergamaschi, P., Krol, M., Meirink, J. F., Dentener, F., Segers, A., van Aardenne, J., Monni, S., Vermeulen, A. T., Schmidt, M., Ramonet, M., Yver, C., Meinhardt, F., Nisbet, E. G., Fisher, R. E., O'Doherty, S., and Dlugokencky, E. J.: Inverse modeling of European CH_4 emissions 2001–2006, Journal of Geophysical Research: Atmospheres, 115, doi:10.1029/2010JD014180, 2010.

Boltzmann, L.: Ableitung des Stefan'schen Gesetzes, betreffend die Abhängigkeit der Wärmestrahlung von der Temperatur aus der electromagnetischen Lichttheorie, Annalen der Physik, 258, 291–294, doi:10.1002/andp.18842580616, URL http://dx.doi.org/10.1002/andp.18842580616, 1884.

Bruhwiler, L. M. P., Michalak, A. M., Peters, W., Baker, D. F., and Tans, P.: An improved Kalman Smoother for atmospheric inversions, Atmospheric Chemistry and Physics, 5, 2691–2702, doi:10.5194/acp-5-2691-2005, URL http://www.atmos-chem-phys.net/5/2691/2005/, 2005.

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

Butz, A., Guerlet, S., Hasekamp, O., Schepers, D., Galli, A., Aben, I., Frankenberg, C., Hartmann, J.-M., Tran, H., Kuze, A., Keppel-Aleks, G., Toon, G., Wunch, D., Wennberg, P., Deutscher, N., Griffith, D., Macatangay, R., Messerschmidt, J., Notholt, J., and Warneke, T.: Toward accurate CO_2 and CH_4 observations from GOSAT, Geophys. Res. Lett., 38, L14812, doi:10.1029/2011GL047888, 2011.

Butz, A., Guerlet, S., Hasekamp, O. P., Kuze, A., and Suto, H.: Using ocean-glint scattered sunlight as a diagnostic tool for satellite remote sensing of greenhouse gases, Atmospheric Measurement Techniques, 6, 2509–2520, doi:10.5194/amt-6-2509-2013, URL http://www.atmos-meas-tech.net/6/2509/2013/, 2013.

Chadwick, R., Good, P., Andrews, T., and Martin, G.: Surface Warming Patterns Drive Tropical Rainfall Pattern Responses to CO2 Forcing on All Timescales, Geophysical Research Letters, doi:10.1002/2013GL058504, 2013.

Chatterjee, A. and Michalak, A. M.: Technical Note: Comparison of ensemble Kalman filter and variational approaches for CO_2 data assimilation, Atmospheric Chemistry and Physics, 13, 11 643–11 660, doi:10.5194/acp-13-11643-2013, URL http://www.atmos-chem-phys.net/13/11643/2013/, 2013.

Chevallier, F.: On the statistical optimality of CO_2 atmospheric inversions assimilating CO_2 column retrievals, Atmospheric Chemistry and Physics Discussions, 15, 11 889–11 923, doi:10.5194/acpd-15-11889-2015, URL http://www.atmos-chem-phys-discuss.net/15/11889/2015/, 2015.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO_2 sources and sinks from satellite observations: Method and application to TOVS data, Journal of Geophysical Research, 110, 24 309, doi:10.1029/2005JD006390, 2005.

Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, Journal of Geophysical Research, 115, doi:10.1029/2010JD013887, 2010.

Chevallier, F., Deutscher, N. M., Conway, T. J., Ciais, P., Ciattaglia, L., Dohe, S., Fröhlich, M., Gomez-Pelaez, A. J., Griffith, D., Hase, F., Haszpra, L., Krummel, P., Kyrö, E., Labuschagne, C., Langenfelds, R., Machida, T., Maignan, F., Matsueda, H., Morino, I., Notholt, J., Ramonet, M., Sawa, Y., Schmidt, M., Sherlock, V., Steele, P., Strong, K., Sussmann, R., Wennberg, P., Wofsy, S., Worthy, D., Wunch, D., and Zimnoch, M.: Global CO2fluxes inferred from surface air-sample measurements and from TCCON retrievals of the CO2 total column, Geophysical Research Letters, 38, doi:10.1029/2011GL049899, 2011.

Cholesky, A.-L.: Sur la résolution numérique des systèmes d'équations linéaires, republished in the Bulletin de la société des amis de la bibliothèque de l'École polytechnique (SABIX), n°39, december 2005, 1910.

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quere, C., Myneni, R., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles, book section 6, p. 465–570, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.015, URL www.climatechange2013.org, 2013.

Ciuraru, R., Fine, L., van Pinxteren, M., D'Anna, B., Herrmann, H., and George, C.: Unravelling New Processes at Interfaces: Photochemical Isoprene Production at the Sea Surface, Environmental Science & Technology, doi:10.1021/acs.est.5b02388, URL http://dx.doi.org/10.1021/acs.est.5b02388, 2015.

Curtis, P. S., Hanson, P. J., Bolstad, P., Barford, C., Randolph, J., Schmid, H., and Wilson, K. B.: Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests, Agricultural and Forest Meteorology, 113, 3–19, doi:10.1016/S0168-1923(02)00099-0, URL http://www.sciencedirect.com/science/article/pii/S0168192302000990, 2002.

Davidson, E., Savage, K., Bolstad, P., Clark, D., Curtis, P., Ellsworth, D., Hanson, P., Law, B., Luo, Y., Pregitzer, K., Randolph, J., and Zak, D.: Belowground carbon allocation in forests estimated from litterfall and IRGA-based soil respiration measurements, Agricultural and Forest Meteorology, 113, 39–51, doi:10.1016/S0168-1923(02)00101-6, URL http://www.sciencedirect.com/science/article/pii/S0168192302001016, 2002.

Davis, S. P., Abrams, M. C., and Brault, J. W.: Fourier Transform Spectrometry, Academic Press, 2001.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L.,

Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, 2011.

Dils, B., Buchwitz, M., Reuter, M., Schneising, O., Boesch, H., Parker, R., Guerlet, S., Aben, I., Blumenstock, T., Burrows, J. P., Butz, A., Deutscher, N. M., Frankenberg, C., Hase, F., Hasekamp, O. P., Heymann, J., De Mazière, M., Notholt, J., Sussmann, R., Warneke, T., Griffith, D., Sherlock, V., and Wunch, D.: The Greenhouse Gas Climate Change Initiative (GHG-CCI): comparative validation of GHG-CCI SCIAMACHY/ENVISAT and TANSO-FTS/GOSAT CO_2 and CH_4 retrieval algorithm products with measurements from the TCCON, Atmospheric Measurement Techniques, 7, 1723–1744, doi:10.5194/amt-7-1723-2014, URL http://www.atmos-meas-tech.net/7/1723/2014/, 2014.

Doney, S. C., Kleypas, J. A., Sarmiento, J. L., and Falkowski, P. G.: The US JGOFS Synthesis and Modeling Project, 2002.

Enting, I. G.: Green's function methods of tracer inversion, vol. 114, pp. 19–31, American Geophysical Union, Washington, DC, 2000.

Errico, R. M.: What is an adjoint model?, Bulletin of the American Meteorological Society, 78, 2577–2591, 1997.

European Centre for Medium-Range Weather Forecasts: ERA Interim, available via the ECMWF Data Server http://apps.ecmwf.int/datasets/, 2013.

Fairbairn, D., Pring, S. R., Lorenc, A. C., and Roulstone, I.: A comparison of 4DVar with ensemble data assimilation methods, Quarterly Journal of the Royal Meteorological Society, doi:10.1002/qj.2135, 2013.

Feist, D. G., Gerbig, C., Geibel, M., Chen, H., Kolle, O., und S. Baum, M. H., Messerschmidt, J., Notholt, J., Palm, M., Petersen, K., Sampson, B., Petri, C., Warneke, T., Blumenstock, T., Hase, F., Pscheidt, I., Rettinger, M., Sussmann, R., Truong, F., Xueref-Remi, I., Katrynski, K., Maser, R., Klaus, C., Schell, D., and Engemann, S.: The IMECC aircraft campaign: validation of total column CO_2, CH_4 and CO measurements over Europe, IMECC, URL http://www.picarro.com/assets/docs/The_IMECC_aircraft_campaign_-_validation_of_total_column_CO2_CH4_and_CO_measurements_over_Europe.pdf, 2010.

Feng, L., Palmer, P. I., Yang, Y., Yantosca, R. M., Kawa, S. R., Paris, J.-D., Matsueda, H., and Machida, T.: Evaluating a 3-D transport model of atmospheric CO_2 using ground-based, aircraft, and space-borne data, Atmospheric Chemistry and Physics, 11, 2789–2803, doi:10.5194/acp-11-2789-2011, URL http://www.atmos-chem-phys.net/11/2789/2011/, 2011.

Feng, L., Palmer, P. I., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R.: Elevated uptake of CO_2 over Europe inferred from GOSAT XCO_2 retrievals: a real phenomenon or an artefact of the analysis?, Atmospheric Chemistry and Physics Discussions, 15, 1989–2011, doi:10.5194/acpd-15-1989-2015, URL http://www.atmos-chem-phys-discuss.net/15/1989/2015/, 2015.

Fisher, M. and Courtier, P.: Estimating the covariance matrices of analysis and forecast error in variational data assimilation, ECMWF Technical Memorandum, Tech. Rep. 220, European Centre for Medium-Range Weather Forecast, 1995.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of Climate Models, book section 9, p. 741–866, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi: 10.1017/CBO9781107415324.020, URL www.climatechange2013.org, 2013.

Foken, T., Aubinet, M., and Leuning, R.: The Eddy Covariance Method, pp. 1–19, Springer, doi:10.1007/978-94-007-2351-1_1, 2012.

Gamnitzer, U., Karstens, U., Kromer, B., Neubert, R. E. M., Meijer, H. A. J., Schroeder, H., and Levin, I.: Carbon monoxide: A quantitative tracer for fossil fuel CO_2?, Journal of Geophysical Research, 111, 22 302, doi:10.1029/2005JD006966, 2006.

Gilbert, J. and Lemaréchal, C.: Some numerical experiments with variable-storage quasi-Newton algorithms, Mathematical Programming, 45, 407–435, doi:10.1007/BF01589113, 1989.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J. M., Brunet, Y., Carrara, A., Clement, R., Dellwik, E., Elbers, J., Eugster, W., Fuhrer, J., Granier, A., Grünwald, T., Heinesch, B., Janssens, I. A., Knohl, A., Koeble, R., Laurila, T., Longdoz, B., Manca, G., Marek, M., Markkanen, T., Mateus, J., Matteucci, G., Mauder, M., Migliavacca, M., Minerbi, S., Moncrieff, J., Montagnani, L., Moors, E., Ourcival, J.-M., Papale, D., Pereira, J., Pilegaard, K., Pita, G., Rambal, S., Rebmann, C., Rodrigues, A., Rotenberg, E., Sanz, M. J., Sedlak, P., Seufert, G., Siebicke, L., Soussana, J. F., Valentini, R., Vesala, T., Verbeeck, H., and Yakir, D.: Quality control of CarboEurope flux data - Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems, Biogeosciences, 5, 433–450, 2008.

Golub, G. and Reinsch, C.: Singular value decomposition and least squares solutions, Numerische Mathematik, 14, 403–420, doi:10.1007/BF02163027, URL http://dx.doi.org/10.1007/BF02163027, 1970.

Griffin, K. L., Anderson, O. R., Gastrich, M. D., Lewis, J. D., Lin, G., Schuster, W., Seemann, J. R., Tissue, D. T., Turnbull, M. H., and Whitehead, D.: Plant growth in elevated CO(2) alters mitochondrial number and chloroplast fine structure, Proc Natl Acad Sci U S A, 98, 2473–2478, doi: 10.1073/pnas.041620898, URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC30162/, 2001.

Griffith, P. R. and de Haseth, J. A.: Fourier Transform Infrared Spectrometry, John Wiley Sons, Inc., 2 edn., 2007.

Gurney, K., Law, R., and Rayner, P.: Transcom 3 experimental protocol, Department of Atmospheric Science, Colorado State University USA, 2000.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, doi:10.1038/415626a, 2002.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Pak, B. C., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyutov, S., Peylin, P., Prather, M., and Taguchi, S.: Transcom 3 inversion intercomparison: Model

mean results for the estimation of seasonal carbon sources and sinks, Global Biogeochemical Cycles, 18, 1010, doi:10.1029/2003GB002111, 2004.

Halliday, D., Resnick, R., Walker, J., and Koch, S.: Physik, Wiley, URL https://books.google.de/books?id=EZ9rAAAACAAJ, 2003.

Hammerling, D. M., Michalak, A. M., and Kawa, S. R.: Mapping of CO2 at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, Journal of Geophysical Research, 117, D06 306, doi:10.1029/2011JD017015, 2012.

Hase, F., Frey, M., Blumenstock, T., Groß, J., Kiel, M., Kohlhepp, R., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Application of portable FTIR spectrometers for detecting greenhouse gas emissions of the major city Berlin, Atmospheric Measurement Techniques, 8, 3059–3068, doi:10.5194/amt-8-3059-2015, URL http://www.atmos-meas-tech.net/8/3059/2015/, 2015.

Hooghiemstra, P. B., Krol, M. C., Meirink, J. F., Bergamaschi, P., van der Werf, G. R., Novelli, P. C., Aben, I., and Röckmann, T.: Optimizing global CO emission estimates using a four-dimensional variational data assimilation system and surface network observations, Atmospheric Chemistry and Physics, 11, 4705–4723, doi:10.5194/acp-11-4705-2011, URL http://www.atmos-chem-phys.net/11/4705/2011/, 2011.

Hortal, M. and Simmons, A. J.: Use of Reduced Gaussian Grids in Spectral Models, Monthly Weather Review, 119, 1057–1074, doi:10.1175/1520-0493(1991)119<1057:UORGGI>2.0.CO;2, 1991.

Houweling, S., Krol, M., Bergamaschi, P., Frankenberg, C., Dlugokencky, E. J., Morino, I., Notholt, J., Sherlock, V., Wunch, D., Beck, V., Gerbig, C., Chen, H., Kort, E. A., Röckmann, T., and Aben, I.: A multi-year methane inversion using SCIAMACHY, accounting for systematic errors using TCCON measurements, Atmospheric Chemistry and Physics, 14, 3991–4012, doi:10.5194/acp-14-3991-2014, URL http://www.atmos-chem-phys.net/14/3991/2014/, 2014.

Höpfner, M., Milz, M., Buehler, S., Orphal, J., and Stiller, G.: The natural greenhouse effect of atmospheric oxygen (O2) and nitrogen (N2), Geophysical Research Letters, 39, n/a–n/a, doi:10.1029/2012GL051409, 2012.

IPCC: Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report, Geneva: IPCC, URL http://www.ipcc.ch/ipccreports/ar4-wg1.htm, 2007.

IPCC: Summary for Policymakers, book section SPM, p. 1–30, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.004, URL www.climatechange2013.org, 2013.

IPCC Working Group I: Climate Change 1990 The Science of Climate Change, The Intergovernmental Panel on Climate Change, 1996a.

IPCC Working Group I: Climate Change 1995 The Science of Climate Change, The Intergovernmental Panel on Climate Change, 1996b.

J. A. Curry, P. J. W.: Climate Science and the Uncertainty Monster, BAMS, pp. 1667–1682, 2011.

Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press, URL http://acmg.seas.harvard.edu/people/faculty/djj/book/, 1999.

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

Jacobson, A. R., Mikaloff Fletcher, S. E., Gruber, N., Sarmiento, J. L., and Gloor, M.: A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes, Global Biogeochemical Cycles, 21, B1019, doi:10.1029/2005GB002556, 2007.

Jones, P. D., New, M., Parker, D. E., Martin, S., and Rigor, I. G.: Surface air temperature and its changes over the past 150 years, Reviews of Geophysics, 37, 173–199, doi:10.1029/1999RG900002, 1999.

Juhász, Z. and Bölöni, G.: Tests with the CONGRAD minimization algorithm within the ALADIN/HU 3DVAR system, URL http://www.rclace.eu/File/Data_Assimilation/2007/jz_bg_congrad.pdf, 2007.

Kalnay, E.: The Future of Data Assimilation: 4D-Var or Ensemble Kalman Filter?, Talk at SAMSI, URL https://web.archive.org/web/20120915073405/http://www.atmos.umd.edu/~ekalnay/4DVarEnKFKalnaySAMSI.ppt.pdf, 2005.

Keeling, C. D., Piper, S. C., Whorf, T. P., and Keeling, R. F.: Evolution of natural and anthropogenic fluxes of atmospheric CO2 from 1957 to 2003, Tellus B, 63, 1–22, doi:10.1111/j.1600-0889.2010.00507.x, 2011.

Khvorostyanov, D. V., Ciais, P., Krinner, G., Zimov, S. A., Corradi, C., and Guggenberger, G.: Vulnerability of permafrost carbon to global warming. Part II: sensitivity of permafrost carbon stock to global warming, Tellus B, 60, 265–275, doi:10.1111/j.1600-0889.2007.00336.x, 2008.

Kiehl, J. T. and Trenberth, K. E.: Earth's Annual Global Mean Energy Budget, Bulletin of the American Meteorological Society, 78, 197–208, doi:10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2, 1997.

Krol, M., Houweling, S., Bregman, B., van den Broek, M., Segers, A., van Velthoven, P., Peters, W., Dentener, F., and Bergamaschi, P.: The two-way nested global chemistry-transport zoom model TM5: algorithm and applications, Atmospheric Chemistry and Physics, 5, 417–432, doi:10.5194/acp-5-417-2005, URL http://www.atmos-chem-phys.net/5/417/2005/, 2005.

Kuang, X., Zhang, Y., Huang, Y., and Huang, D.: Changes in the Frequencies of Record-breaking Temperature Events in China and Its Association with East Asian Winter Monsoon Variability, Journal of Geophysical Research: Atmospheres, doi:10.1002/2013JD020965, 2014.

Lahoz, W. A. and Schneider, P.: Data Assimilation: Making Sense of Earth Observation, Frontiers in Environmental Science, 2, doi:10.3389/fenvs.2014.00016, 2014.

Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D.: Data assimilation of stratospheric constituents: a review, Atmospheric Chemistry and Physics, 7, 5745–5773, doi:10.5194/acp-7-5745-2007, URL http://www.atmos-chem-phys.net/7/5745/2007/, 2007.

Lanczos, C.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, Journal of Research of the National Bureau of Standards, 45, 255–282, 1950.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, Earth System Science Data Discussions, 5, 1107–1157, doi:10.5194/essdd-5-1107-2012, URL http://www.earth-syst-sci-data-discuss.net/5/1107/2012/, 2012.

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., Friedlingstein, P., Jones, S. D., Sitch, S., Tans, P., Arneth, A., Boden, T. A., Bopp, L., Bozec, Y., Canadell, J. G., Chini, L. P., Chevallier, F., Cosca, C. E., Harris, I., Hoppema, M., Houghton, R. A., House, J. I., Jain, A. K., Johannessen, T., Kato, E., Keeling, R. F., Kitidis, V., Klein Goldewijk, K., Koven, C., Landa, C. S., Landschützer, P., Lenton, A., Lima, I. D., Marland, G., Mathis, J. T., Metzl, N., Nojiri, Y., Olsen, A., Ono, T., Peng, S., Peters, W., Pfeil, B., Poulter, B., Raupach, M. R., Regnier, P., Rödenbeck, C., Saito, S., Salisbury, J. E., Schuster, U., Schwinger, J., Séférian, R., Segschneider, J., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Werf, G. R., Viovy, N., Wang, Y.-P., Wanninkhof, R., Wiltshire, A., and Zeng, N.: Global carbon budget 2014, Earth System Science Data, 7, 47–85, doi:10.5194/essd-7-47-2015, URL http://www.earth-syst-sci-data.net/7/47/2015/, 2015.

Luyssaert, S., Inglima, I., and Jung, M.: Global Forest Ecosystem Structure and Function Data for Carbon Balance Research. Data set. Available on-line [http://daac.ornl.gov/] from Oak Ridge National Laboratory, Distributed Active Archive Center, Oak Ridge, Tennessee, U.S.A., doi:10.3334/ORNLDAAC/949, 2009.

Machida, T., Matsueda, H., Sawa, Y., Nakagawa, Y., Hirotani, K., Kondo, N., Goto, K., Nakazawa, T., Ishikawa, K., and Ogawa, T.: Worldwide Measurements of Atmospheric CO2 and Other Trace Gas Species Using Commercial Airlines, Journal of Atmospheric and Oceanic Technology, 25, 1744–1754, doi:10.1175/2008JTECHA1082.1, URL http://dx.doi.org/10.1175/2008JTECHA1082.1, 2008.

Masarie, K. A., Peters, W., Jacobson, A. R., and Tans, P. P.: ObsPack: a framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements, Earth System Science Data, 6, 375–384, doi:10.5194/essd-6-375-2014, URL http://www.earth-syst-sci-data.net/6/375/2014/, 2014.

McGill, R., Tukey, J. W., and Larsen, W. A.: Variations of Box Plots, The American Statician, 32, 12–16, 1978.

Meesters, A. G. C. A., Tolk, L. F., Peters, W., Hutjes, R. W. A., Vellinga, O. S., Elbers, J. A., Vermeulen, A. T., van der Laan, S., Neubert, R. E. M., Meijer, H. A. J., and Dolman, A. J.: Inverse carbon dioxide flux estimates for the Netherlands, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2012JD017797, 2012.

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmospheric Chemistry and Physics, 8, 6341–6353, doi:10.5194/acp-8-6341-2008, URL http://www.atmos-chem-phys.net/8/6341/2008/, 2008.

Messerschmidt, J., Geibel, M. C., Blumenstock, T., Chen, H., Deutscher, N. M., Engel, A., Feist, D. G., Gerbig, C., Gisi, M., Hase, F., Katrynski, K., Kolle, O., Lavrič, J. V., Notholt, J.,

Palm, M., Ramonet, M., Rettinger, M., Schmidt, M., Sussmann, R., Toon, G. C., Truong, F., Warneke, T., Wennberg, P. O., Wunch, D., and Xueref-Remy, I.: Calibration of TCCON column-averaged CO_2: the first aircraft campaign over European TCCON sites, Atmospheric Chemistry and Physics, 11, 10 765–10 777, doi:10.5194/acp-11-10765-2011, URL http://www.atmos-chem-phys.net/11/10765/2011/, 2011.

Myhre, G., Shindell, D., Bre'on, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing, book section 8, p. 659–740, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.018, URL www.climatechange2013.org, 2013.

Navon, I. M. and Legler, D. M.: Conjugate-Gradient Methods for Large-Scale Minimization, in Meteorology, Monthly Weather Review, 115, 1479–1502, 1987.

Newton, S. I.: De analysi per æquationes numero terminorum infinitas, William Jones, 1669.

NOAA Environmental Sciences Division, Oak Ridge National Laboratory: obspack CO_2 1 PROTOTYPE v1.0.4 2013-11-25, doi:10.3334/OBSPACK/1001, 2013.

NOAA Global Monitoring Division: Boulder, Colorado, U.S.A.: Cooperative Global Atmospheric Data Integration Project. 2013, updated annually. Multi-laboratory compilation of synchronized and gap-filled atmospheric carbon dioxide records for the period 1979-2012 (obspack co2 1 PROTOTYPE v1.0.2 2013-01-28), doi:10.3334/OBSPACK/1002, 2013.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Eldering, D., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO_2 retrieval algorithm – Part 1: Description and validation against synthetic observations, Atmospheric Measurement Techniques, 5, 99–121, doi:10.5194/amt-5-99-2012, URL http://www.atmos-meas-tech.net/5/99/2012/, 2012.

Olson, J., Watts, J., and Allsion, L.: Major World Ecosystem Complexes Ranked by Carbon in Live Vegetation: A Database, Tech. rep., Carbon Dioxide Information Analysis Center, U.S. Department of Energy, Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A, doi:10.3334/CDIAC/lue.ndp017, URL http://cdiac.ornl.gov/epubs/ndp/ndp017/ndp017b.html, 1992.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and Yakir, D.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571–583, 2006.

Parker, R., Boesch, H., Cogan, A., Fraser, A., Feng, L., Palmer, P. I., Messerschmidt, J., Deutscher, N., Griffith, D. W. T., Notholt, J., Wennberg, P. O., and Wunch, D.: Methane observations from the Greenhouse Gases Observing SATellite: Comparison to ground-based TCCON data and model calculations, Geophysical Research Letters, 38, n/a–n/a, doi:10.1029/2011GL047871, URL http://dx.doi.org/10.1029/2011GL047871, 2011.

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

Y., Onishi, R., Parazoo, N., Pieterse, G., Rivier, L., Satoh, M., Serrar, S., Taguchi, S., Vautard, R., Vermeulen, A. T., and Zhu, Z.: TransCom model simulations of hourly atmospheric CO2: Analysis of synoptic-scale variations for the period 2002–2003, Global Biogeochemical Cycles, 22, n/a–n/a, doi:10.1029/2007GB003081, URL http://dx.doi.org/10.1029/2007GB003081, 2008.

Peters, G. P., Andrew, R. M., Boden, T., Canadell, J. G., Ciais, P., Le Quere, C., Marland, G., Raupach, M. R., and Wilson, C.: The challenge to keep global warming below 2°C, Nature Climate Change, 3, 4–6, doi:10.1038/nclimate1783, 2013.

Peters, W.: SiBCASA, description, personal communication, 2014.

Peters, W., Miller, J. B., Whitaker, J., Denning, A. S., Hirsch, A., Krol, M. C., Zupanski, D., Bruhwiler, L., and Tans, P. P.: An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations, Journal of Geophysical Research: Atmospheres, 110, 24 304, doi:10.1029/2005JD006157, 2005.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, Proceedings of the National Academy of Sciences, 104, 18 925–18 930, doi:10.1073/pnas.0708986104, URL http://www.pnas.org/content/104/48/18925.abstract, 2007.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Heltai, D., Meinhardt, F., di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Ström, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, A., Rodó, X., Morguí, J.-A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, Global Change Biology, 16, 1317–1337, doi:10.1111/j.1365-2486.2009.02078.x, 2010.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., van der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: results from an ensemble of atmospheric CO_2 inversions, Biogeosciences, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, URL http://www.biogeosciences.net/10/6699/2013/, 2013.

Portnov, A., Smith, A. J., Mienert, J., Cherkashov, G., Rekant, P., Semenov, P., Serov, P., and Vanshtein, B.: Offshore permafrost decay and massive seabed methane escape in water depths >20m at the South Kara Sea shelf, Geophysical Research Letters, doi:10.1002/grl.50735, 2013.

Ramanathan, V. and Coakley, J. A.: Climate modeling through radiative-convective models, Reviews of Geophysics, 16, 465–489, doi:10.1029/RG016i004p00465, URL http://www.phy.pku.edu.cn/climate/class/cm2010/Ramanathan-1978.pdf, 1978.

Rayner, N. A., Brohan, P., Parker, D. E., Folland, C. K., Kennedy, J. J., Vanicek, M., Ansell, T. J., and Tett, S. F. B.: Improved Analyses of Changes and Uncertainties in Sea Surface Temperature Measured In Situ since the Mid-Nineteenth Century: The HadSST2 Dataset, Journal of Climate, 19, 446–469, doi:10.1175/JCLI3637.1, 2006.

Reuter, M., Bovensmann, H., Buchwitz, M., Burrows, J. P. Connor, B. J., Deutscher, N. M., Griffith, D. W. T., Heymann, J., Keppel-Aleks, G., Messerschmidt, J., Notholt, J., Petri, C., Robinson, J., Schneising, O., Sherlock, V., Velazco, V., Warneke, T., Wennberg, P. O., and Wunch, D.: Retrieval of atmospheric CO_2 with enhanced accuracy and precision from SCIAMACHY: Validation with FTS measurements and comparison with model results., JGR, 116, doi:10.1029/2010JD015047, 2011.

Reuter, M., Buchwitz, M., Hilker, M., Heymann, J., Schneising, O., Pillai, D., Bovensmann, H., Burrows, J. P., Bösch, H., Parker, R., Butz, A., Hasekamp, O., O'Dell, C. W., Yoshida, Y., Gerbig, C., Nehrkorn, T., Deutscher, N. M., Warneke, T., Notholt, J., Hase, F., Kivi, R., Sussmann, R., Machida, T., Matsueda, H., and Sawa, Y.: Satellite-inferred European carbon sink larger than expected, Atmospheric Chemistry and Physics, 14, 13 739–13 753, doi:10.5194/acp-14-13739-2014, URL http://www.atmos-chem-phys.net/14/13739/2014/, 2014.

Rhein, M., Rintoul, S., Aoki, S., Campos, E., Chambers, D., Feely, R., Gulev, S., Johnson, G., Josey, S., Kostianoy, A., Mauritzen, C., Roemmich, D., Talley, L., and Wang, F.: Observations: Ocean, book section 3, p. 255–316, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.010, URL www.climatechange2013.org, 2013.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding : Theory and Practice (Series on Atmospheric Oceanic and Planetary Physics), World Scientific Publishing Company, URL http://www.worldcat.org/isbn/981022740X, 2000.

Russell, G. L. and Lerner, J. A.: A New Finite-Differencing Scheme for the Tracer Transport Equation, Journal of Applied Meteorology, 20, "1483–1498", doi:10.1175/1520-0450(1981)020<1483:ANFDSF>2.0.CO;2, 1981.

Sabine, C. L., Key, R. M., Kozyr, A., Feely, R. A., Wanninkhof, R., Millero, F. J., Peng, T. H., Bullister, J. L., and Lee., K.: Global Ocean Data Analysis Project: Results and Data, Tech. rep., Carbon Dioxide Information Analysis Center Oak Ridge National Laboratory Oak Ridge, Tennessee, U.S.A., doi:10.3334/CDIAC/otg.ndp083, 2005.

Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., Prihodko, L., Suits, N., and Philpott, A.: Combined Simple Biosphere/Carnegie-Ames-Stanford Approach terrestrial carbon cycle model, Journal of Geophysical Research: Biogeosciences, 113, doi:10.1029/2007JG000603, 2008.

Schmidt, G. A., Ruedy, R. A., Miller, R. L., and Lacis, A. A.: Attribution of the present-day total greenhouse effect, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2010JD014287, URL http://dx.doi.org/10.1029/2010JD014287, 2010.

Schneider von Deimling, T., Meinshausen, M., Levermann, A., Huber, V., Frieler, K., Lawrence, D. M., and Brovkin, V.: Estimating the near-surface permafrost-carbon feedback on global warming, Biogeosciences, 9, 649–665, doi:10.5194/bg-9-649-2012, URL http://www.biogeosciences.net/9/649/2012/, 2012.

Schulze, E.-D.: Biological control of the terrestrial carbon sink, Biogeosciences, 3, 147–166, doi:10.5194/bg-3-147-2006, URL http://www.biogeosciences.net/3/147/2006/, 2006.

Schulze, E. D., Luyssaert, S., Ciais, P., Freibauer, A., Janssens, I. A., Soussana, J. F., Smith, P., Grace, J., Levin, I., Thiruchittampalam, B., Heimann, M., Dolman, A. J., Valentini, R.,

Bousquet, P., Peylin, P., Peters, W., Rödenbeck, C., Etiope, G., Vuichard, N., Wattenbach, M., Nabuurs, G. J., Poussi, Z., Nieschulze, J., Gash, J. H., and the CarboEurope Team: Importance of methane and nitrous oxide for Europe's terrestrial greenhouse-gas balance, Nature Geoscience, 2, 842–850, doi:10.1038/ngeo686, 2009.

Skachko, S., Errera, Q., Ménard, R., Christophe, Y., and Chabrillat, S.: Comparison of the ensemble Kalman filter and 4D-Var assimilation methods using a stratospheric tracer transport model, Geoscientific Model Development, 7, 1451–1465, doi:10.5194/gmd-7-1451-2014, URL http://www.geosci-model-dev.net/7/1451/2014/, 2014.

Sofiev, M., Berger, U., Prank, M., Vira, J., Arteta, J., Belmonte, J., Bergmann, K.-C., Chéroux, F., Elbern, H., Friese, E., Galan, C., Gehrig, R., Khvorostyanov, D., Kranenburg, R., Kumar, U., Marécal, V., Meleux, F., Menut, L., Pessi, A.-M., Robertson, L., Ritenberga, O., Rodinkova, V., Saarto, A., Segers, A., Severova, E., Sauliene, I., Siljamo, P., Steensen, B. M., Teinemaa, E., Thibaudon, M., and Peuch, V.-H.: MACC regional multi-model ensemble simulations of birch pollen dispersion in Europe, Atmospheric Chemistry and Physics, 15, 8115–8130, doi:10.5194/acp-15-8115-2015, URL http://www.atmos-chem-phys.net/15/8115/2015/, 2015.

Stocker, T. F., Dahe, Q., and Plattner, G.-K.: Climate Change 2013: The Physical Science Basis, Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers (IPCC, 2013), 2013.

Tange, O.: GNU Parallel - The Command-Line Power Tool, ;login: The USENIX Magazine, 36, 42–47, URL http://www.gnu.org/s/parallel, 2011.

Thompson, R. L., Ishijima, K., Saikawa, E., Corazza, M., Karstens, U., Patra, P. K., Bergamaschi, P., Chevallier, F., Dlugokencky, E., Prinn, R. G., Weiss, R. F., O'Doherty, S., Fraser, P. J., Steele, L. P., Krummel, P. B., Vermeulen, A., Tohjima, Y., Jordan, A., Haszpra, L., Steinbacher, M., Van der Laan, S., Aalto, T., Meinhardt, F., Popa, M. E., Moncrieff, J., and Bousquet, P.: TransCom N_2O model inter-comparison – Part 2: Atmospheric inversion estimates of N_2O emissions, Atmospheric Chemistry and Physics, 14, 6177–6194, doi:10.5194/acp-14-6177-2014, URL http://www.atmos-chem-phys.net/14/6177/2014/, 2014.

Toon, G., Blavier, J.-F., Washenfelder, R., Wunch, D., Keppel-Aleks, G., Wennberg, P., Connor, B., Sherlock, V., Griffith, D., Deutscher, N., and Notholt, J.: Total Column Carbon Observing Network (TCCON), in: Advances in Imaging, p. JMA3, Optical Society of America, doi:10.1364/FTS.2009.JMA3, URL http://www.opticsinfobase.org/abstract.cfm?URI=FTS-2009-JMA3, 2009a.

Toon, G., Blavier, J.-F., Washenfelder, R., Wunch, D., Keppel-Aleks, G., Wennberg, P., Connor, B., Sherlock, V., Griffith, D., Deutscher, N., and Notholt, J.: Total Column Carbon Observing Network (TCCON), in: Advances in Imaging, p. JMA3, Optical Society of America, doi:10.1364/FTS.2009.JMA3, URL http://www.opticsinfobase.org/abstract.cfm?URI=FTS-2009-JMA3, 2009b.

Trenberth, K. E. and Smith, L.: The Mass of the Atmosphere: A Constraint on Global Analyses, Journal of Climate, 18, 864–875, doi:10.1175/JCLI-3299.1, URL http://dx.doi.org/10.1175/JCLI-3299.1, 2005.

Tschakert, P.: 1.5°C or 2°C: a conduit's view from the science-policy interface at COP20 in Lima, Peru, Climate Change Responses, doi:doi:10.1186/s40665-015-0010-z, URL http://www.climatechangeresponses.com/content/2/1/3, 2015.

Turner, A. J. and Jacob, D. J.: Balancing aggregation and smoothing errors in inverse models, Atmospheric Chemistry and Physics Discussions, 15, 1001–1026, doi:10.5194/acpd-15-1001-2015, URL http://www.atmos-chem-phys-discuss.net/15/1001/2015/, 2015.

U.S. Department of Energy Office of Science, ed.: Carbon Cycling and Biosequestration: Report from the March 2008 Workshop, DOE/SC-108, U.S. Department of Energy Office of Science, URL http://genomicscience.energy.gov/carboncycle/report/CarbonCycle012609HR.pdf, 2008.

van der Velde, I. R., Miller, J. B., Schaefer, K., van der Werf, G. R., Krol, M. C., and Peters, W.: Terrestrial cycling of $^{13}CO_2$ by photosynthesis, respiration, and biomass burning in SiBCASA, Biogeosciences, 11, 6553–6571, doi:10.5194/bg-11-6553-2014, URL http://www.biogeosciences.net/11/6553/2014/, 2014.

van der Werf, G. R., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and van Leeuwen, T. T.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), Atmospheric Chemistry and Physics Discussions, 10, 16 153–16 230, doi:10.5194/acpd-10-16153-2010, URL http://www.atmos-chem-phys-discuss.net/10/16153/2010/, 2010.

Whitaker, J. S. and Hamill, T. M.: Ensemble Data Assimilation without Perturbed Observations, Monthly Weather Review, 130, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2, 2002.

WOCE Data Products Committee: WOCE Global Data, Version 3.0, WOCE Report No. 180/02, Tech. rep., WOCE International Project Office, Southampton, UK, URL http://www.nodc.noaa.gov/woce/wdiu/index.htm, 2002.

Wofsy, S. C.: HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 369, 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

Wolkovich, E. M., Cook, B. I., Allen, J. M., Crimmins, T. M., Betancourt, J. L., Travers, S. E., Pau, S., Regetz, J., Davies, T. J., Kraft, N. J. B., Ault, T. R., Bolmgren, K., Mazer, S. J., McCabe, G. J., McGill, B. J., Parmesan, C., Salamin, N., Schwartz, M. D., and Cleland, E. E.: Warming experiments underpredict plant phenological responses to climate change, Nature, advance online publication, doi:10.1038/nature11014, 2012.

Worrell, E., Price, L., Martin, N., Hendriks, C., and Meida, L. O.: CARBON DIOXIDE EMISSIONS FROM THE GLOBAL CEMENT INDUSTRY, Annual Review of Energy and the Environment, 26, 303–329, doi:10.1146/annurev.energy.26.1.303, URL http://dx.doi.org/10.1146/annurev.energy.26.1.303, 2001.

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

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Rettinger, M., Robinson, J., Roehl, C. M.,

Salawitch, R. J., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R., Uchino, O., Warneke, T., and Wofsy, S. C.: A method for evaluating bias in global measurements of CO_2 total columns from space, Atmospheric Chemistry and Physics, 11, 12 317–12 337, doi:10.5194/acp-11-12317-2011, 2011b.

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

Yoshida, Y., Kikuchi, N., Morino, I., Uchino, O., Oshchepkov, S., Bril, A., Saeki, T., Schutgens, N., Toon, G. C., Wunch, D., Roehl, C. M., Wennberg, P. O., Griffith, D. W. T., Deutscher, N. M., Warneke, T., Notholt, J., Robinson, J., Sherlock, V., Connor, B., Rettinger, M., Sussmann, R., Ahonen, P., Heikkinen, P., Kyrö, E., and Yokota, T.: Improvement of the retrieval algorithm for GOSAT SWIR XCO_2 and XCH_4 and their validation using TCCON data, Atmospheric Measurement Techniques Discussion, 6, 949–988, doi:10.5194/amtd-6-949-2013, 2013.

Zhang, H. F., Chen, B. Z., van der Laan-Luijk, I. T., Machida, T., Matsueda, H., Sawa, Y., Fukuyama, Y., Langenfelds, R., Labuschagne, C., van der Schoot, M., Xu, G., Yan, J. W., Cheng, M. L., Zhou, L. X., Tans, P. P., and Peters, W.: Estimating Asian terrestrial carbon fluxes from CONTRAIL aircraft and surface CO_2 observations for the period 2006–2010, Atmospheric Chemistry and Physics, 14, 5807–5824, doi:10.5194/acp-14-5807-2014, URL http://www.atmos-chem-phys.net/14/5807/2014/, 2014.

Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, Journal of Geophysical Research-Atmospheres, doi:doi:10.1029/2005JD006003, 2006.

Zhao, C. L., Tans, P. P., and Thoning, K. W.: A high precision manometric system for absolute calibrations of CO2 in dry air, Journal of Geophysical Research-Atmospheres, 1997.

Štefan, J.: Über die Beziehung zwischen der Wärmestrahlung und der Temperatur. In: Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften, Sitzungsberichte der mathematisch-naturwissenschaftlichen Classe der kaiserlichen Akademie der Wissenschaften, 79, 391–428, 1879.

Appendix A

Appendix

A.1 Prior flux uncertainty harmonization

We use the flux uncertainties of a Carbontracker run with a monthly instead of weekly cycle, which allows us to avoid the unspecified temporal aggregation of covariances in Carbontracker, and a run of TM5-4DVar as a proxy to harmonize the covariance matrices of the models.

Due to the different specification of fluxes, harmonizing global and regional flux uncertainties is at odds. To understand the impact of the global harmonization on the flux in the different transcom regions (the smallest unit in which we can compare differences between the models), we compare the time series of prior fluxes and flux uncertainties for each region.

The time series in Fig. A.1 shows that there is a mismatch in the flux uncertainty definition, but this mismatch occurs in April and May 2009 and in Winter 2009/2010, while as figure 9 shows, the flux difference in TM5-4DVar from assimilating the site in Arembepe extends from April to August. April and May are the months, however, where the outlier rejection of Carbontracker makes a difference in the estimated weekly.

The prior flux time series of other regions are shown in figures A.2, A.3, A.4, A.5, A.6, A.7, A.8, A.9, A.10, A.11, A.12, A.13, A.14, A.15, A.16, A.17, A.18, A.19, A.20, A.21 and A.22.

A.2 In-situ measurement sites

Table A.1 shows the measurement sites which are either assimilated in the different setups of CarbonTracker and TM5-4DVar or used in the resampling study. Figurse 3.1 and 3.14 show a graphical representation of the observations used in the assimilation and used for validation in the different runs. The sites are referenced by their NOAA sitecode, as defined in OBSPACK from Masarie et al. (2014), compiled by the NOAA Environmental Sciences Division, Oak Ridge National Laboratory (2013, exact version: obspack PROTOTYPE v1.0.2 2013-01-28).

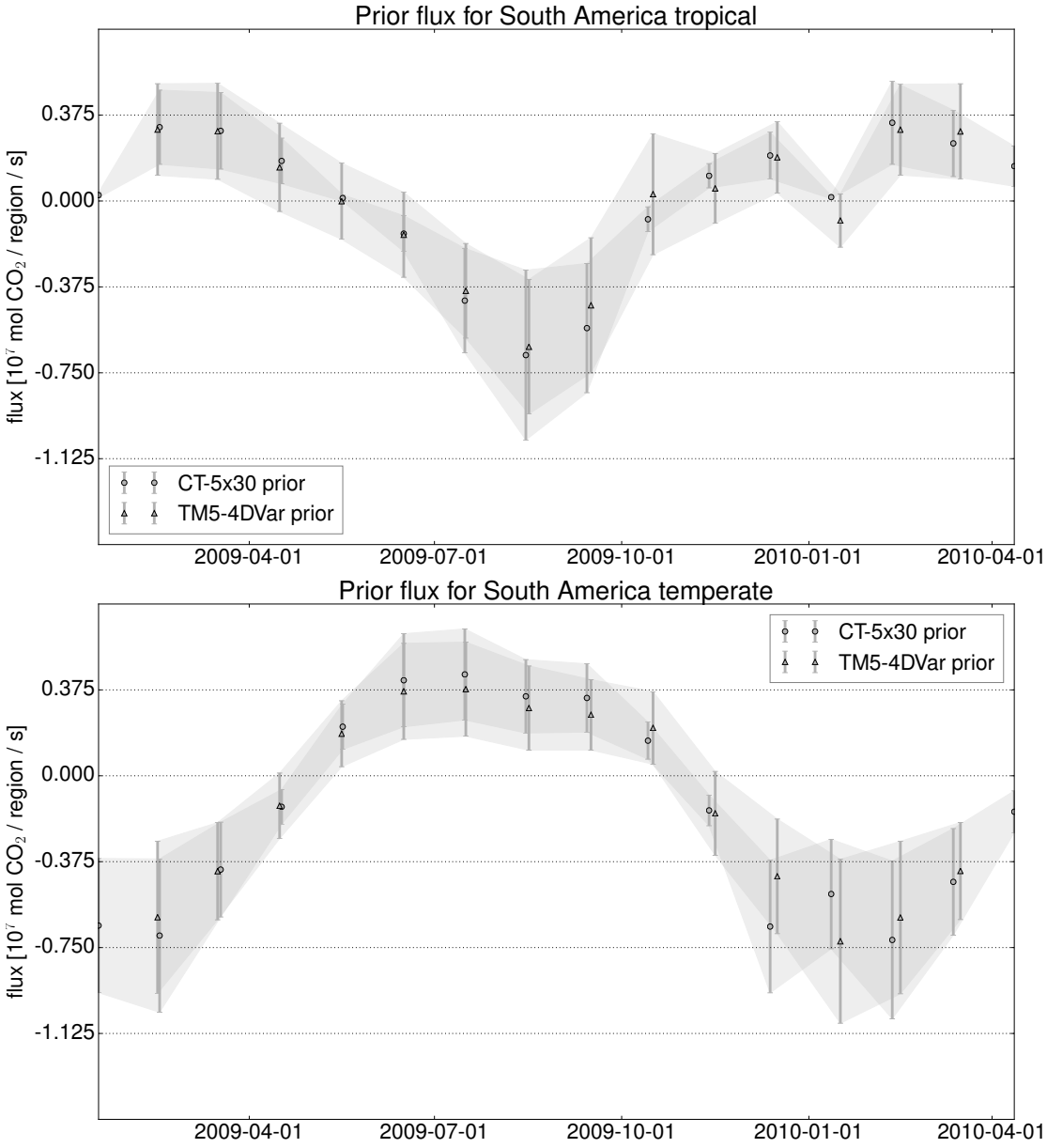


Figure A.1: Monthly prior flux and uncertainty for South America.

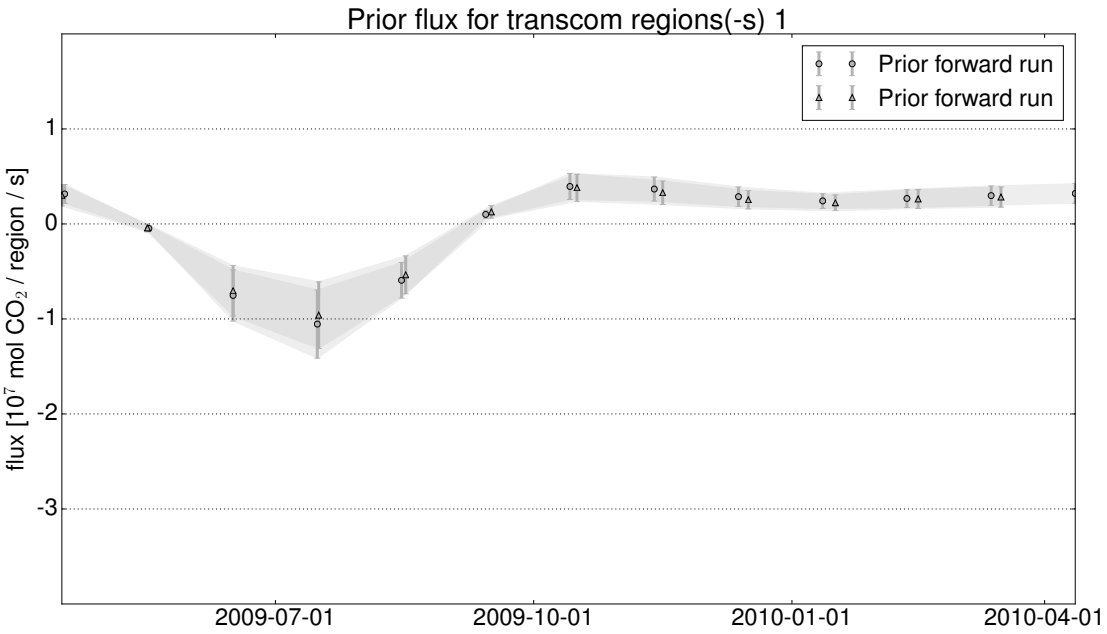


Figure A.2: Monthly prior flux and uncertainty for Boreal North America.

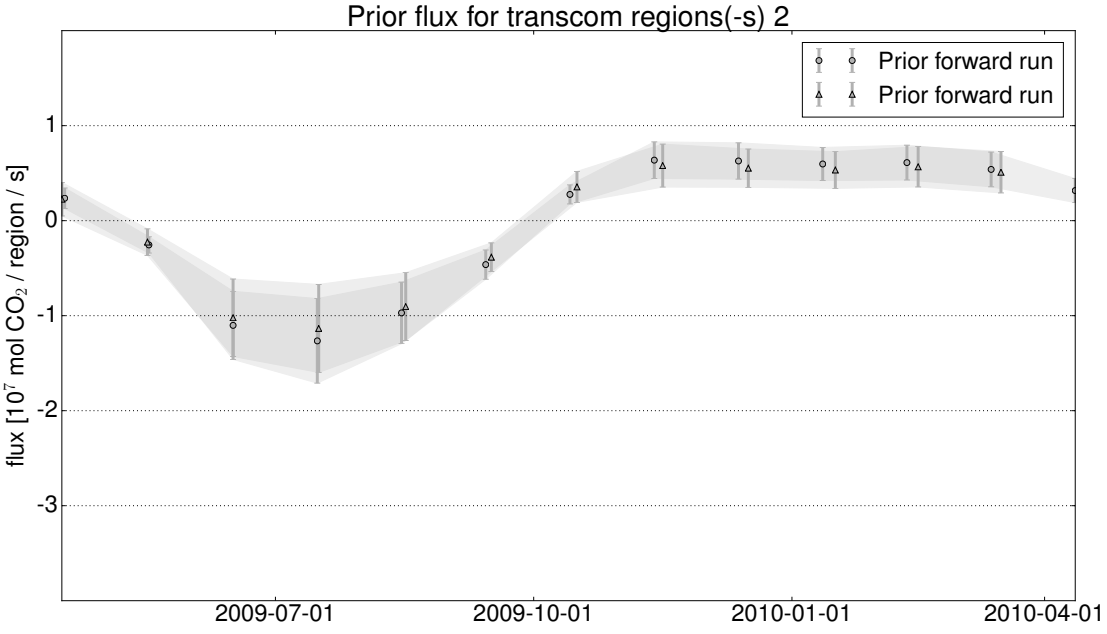


Figure A.3: Monthly prior flux and uncertainty for Temperate North America.

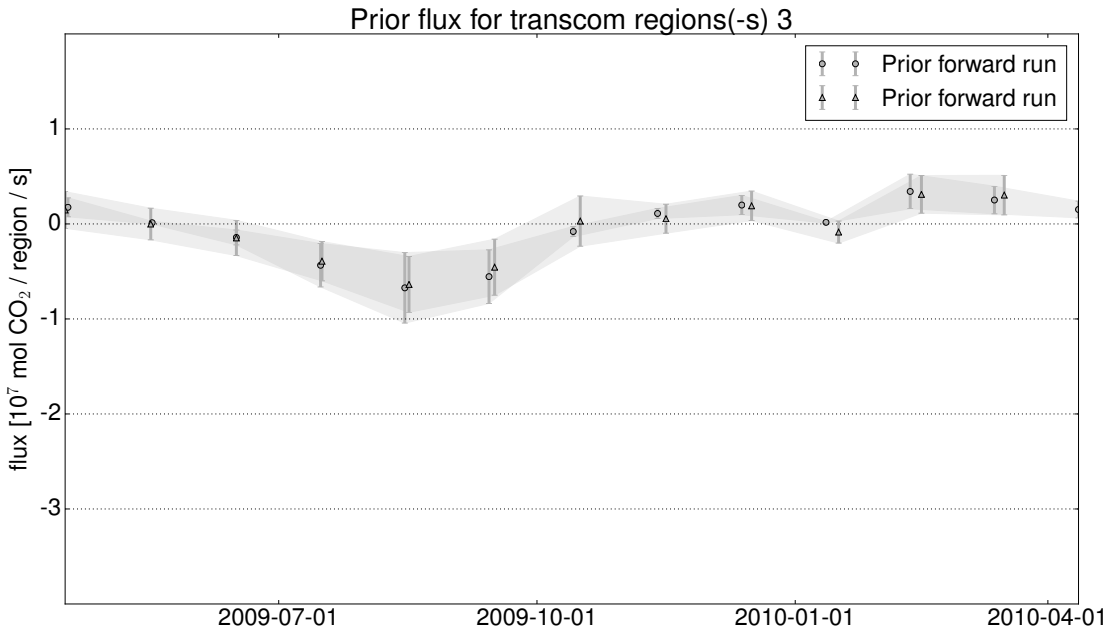


Figure A.4: Monthly prior flux and uncertainty for Tropical South America.

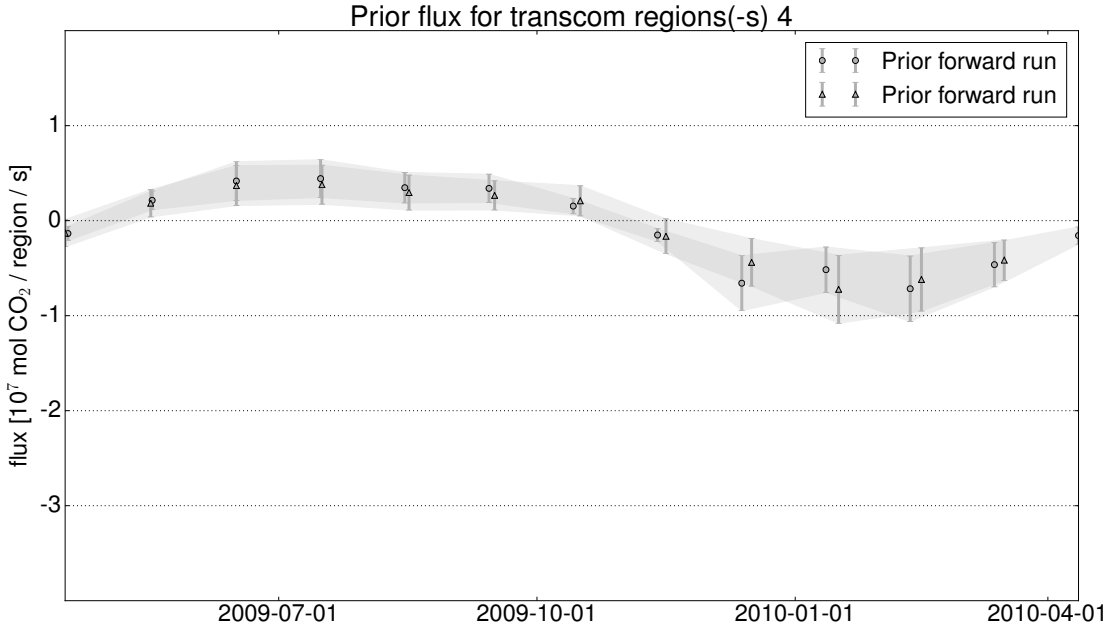


Figure A.5: Monthly prior flux and uncertainty for Temperate South America.

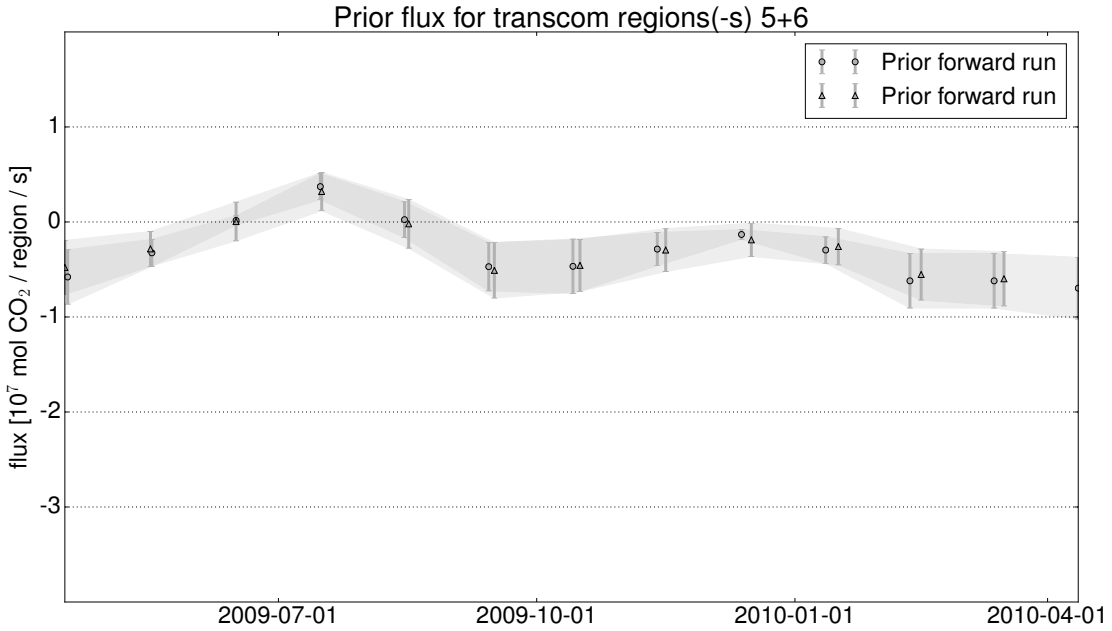


Figure A.6: Monthly prior flux and uncertainty for Africa (both regions combined due to different region specification).

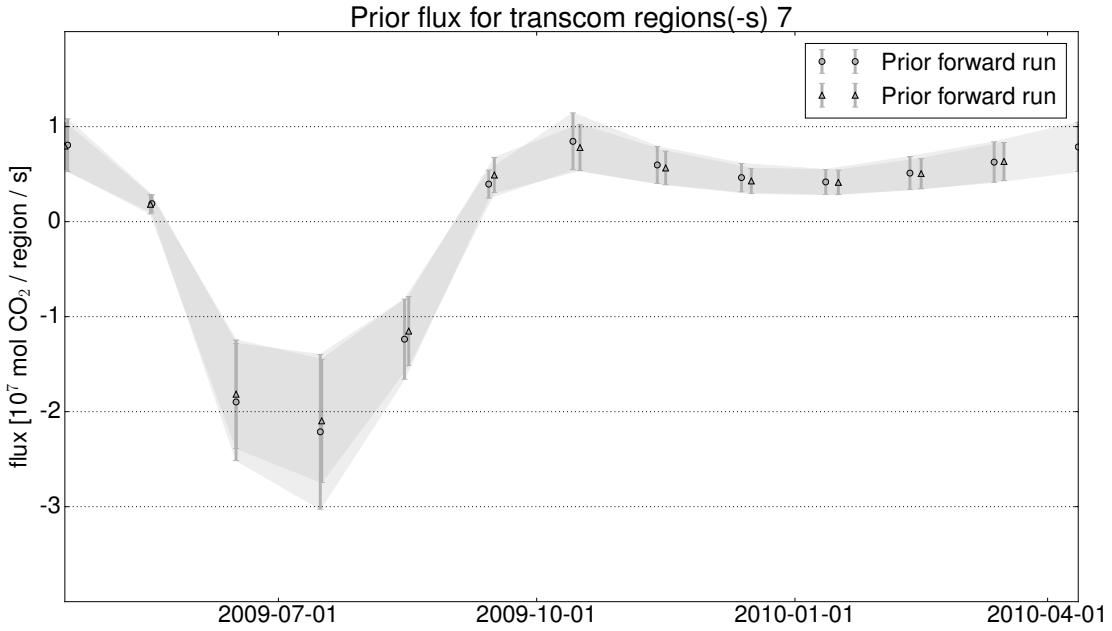


Figure A.7: Monthly prior flux and uncertainty for Boreal Eurasia.

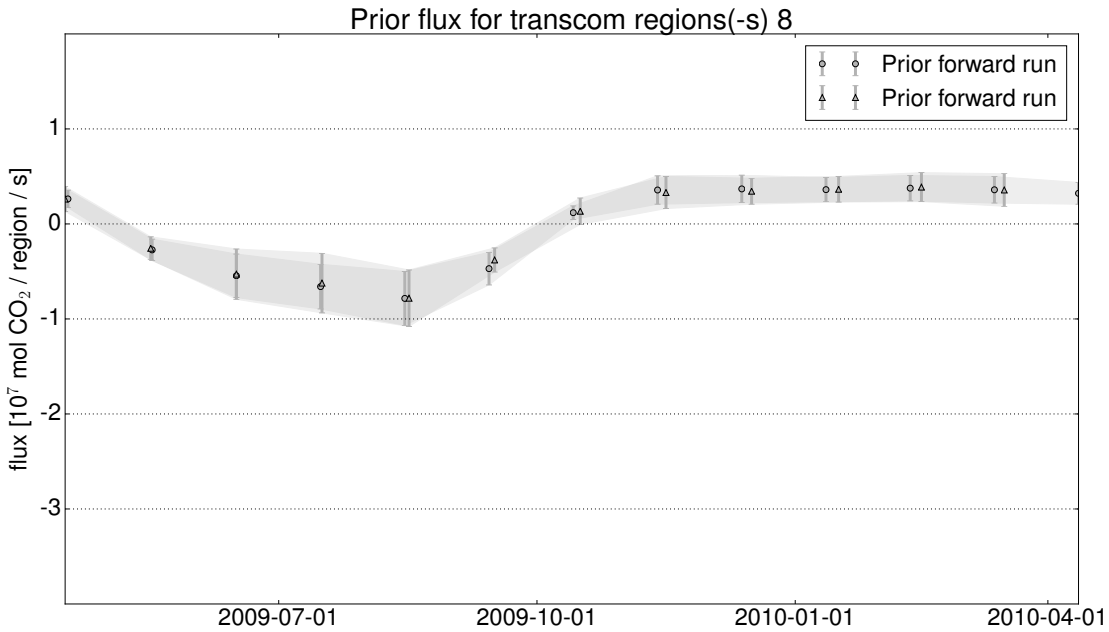


Figure A.8: Monthly prior flux and uncertainty Temperate Eurasia.

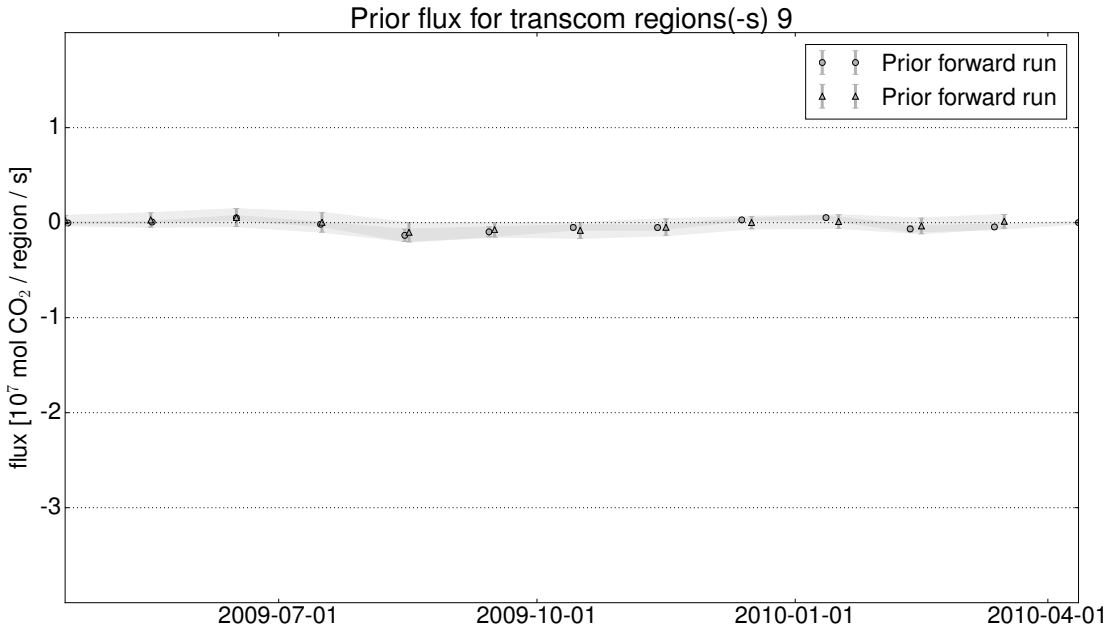


Figure A.9: Monthly prior flux and uncertainty for Tropical Asia.

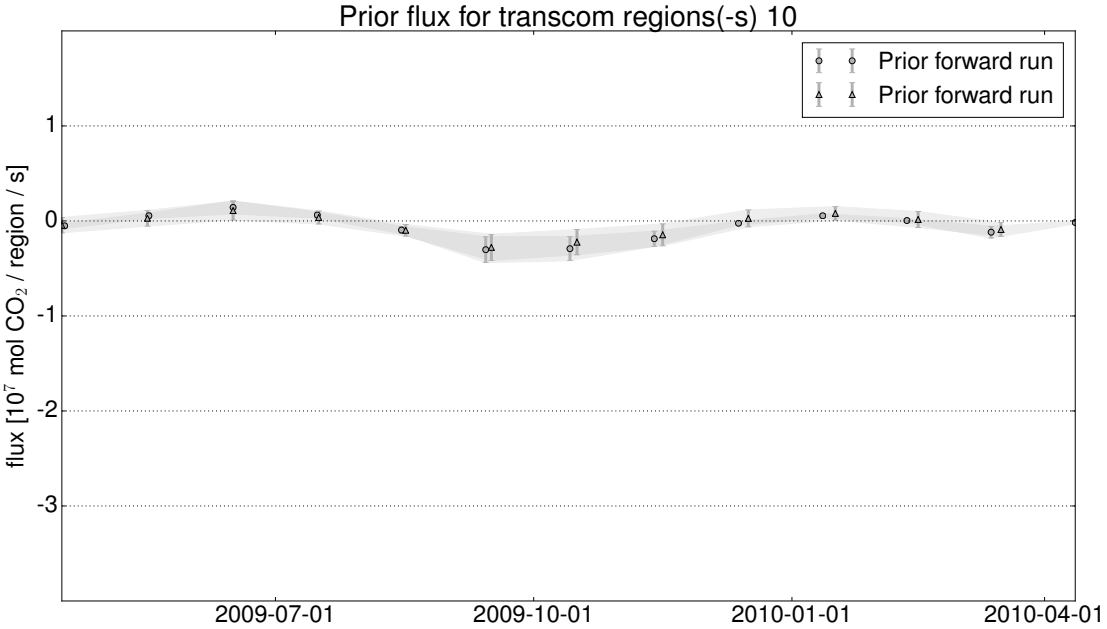


Figure A.10: Monthly prior flux and uncertainty for Australia.

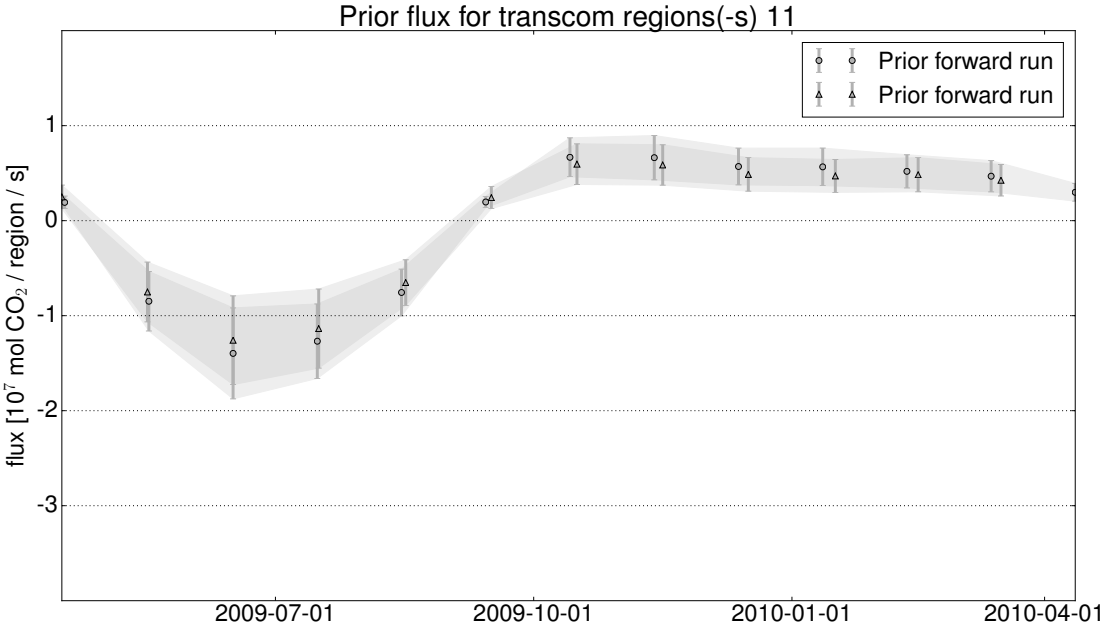


Figure A.11: Monthly prior flux and uncertainty for Europe.

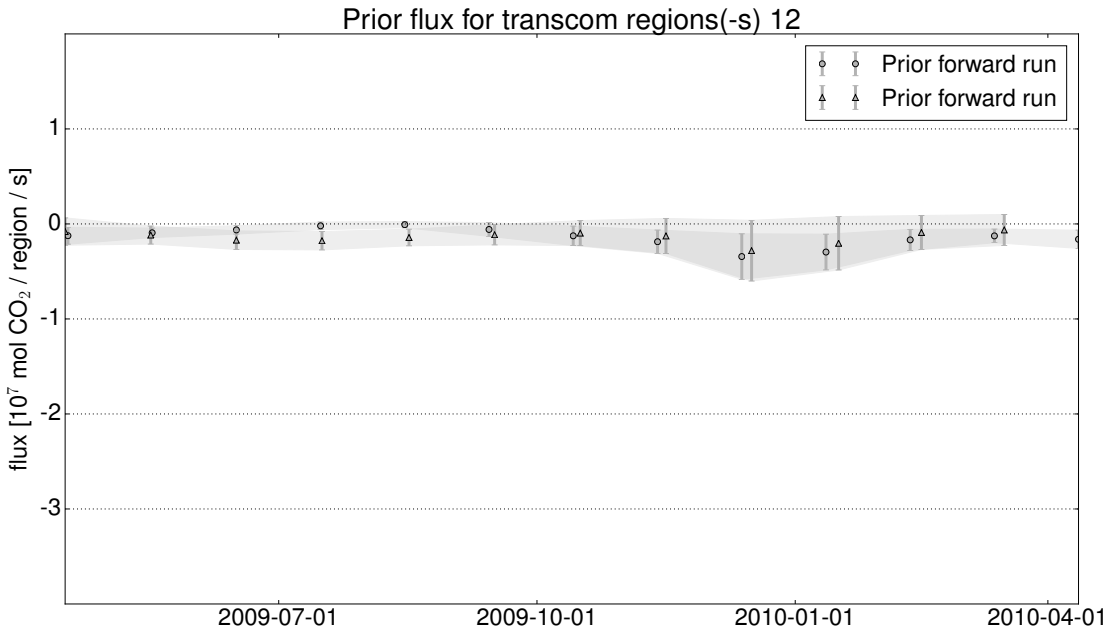


Figure A.12: Monthly prior flux and uncertainty for Temperate North Pacific.

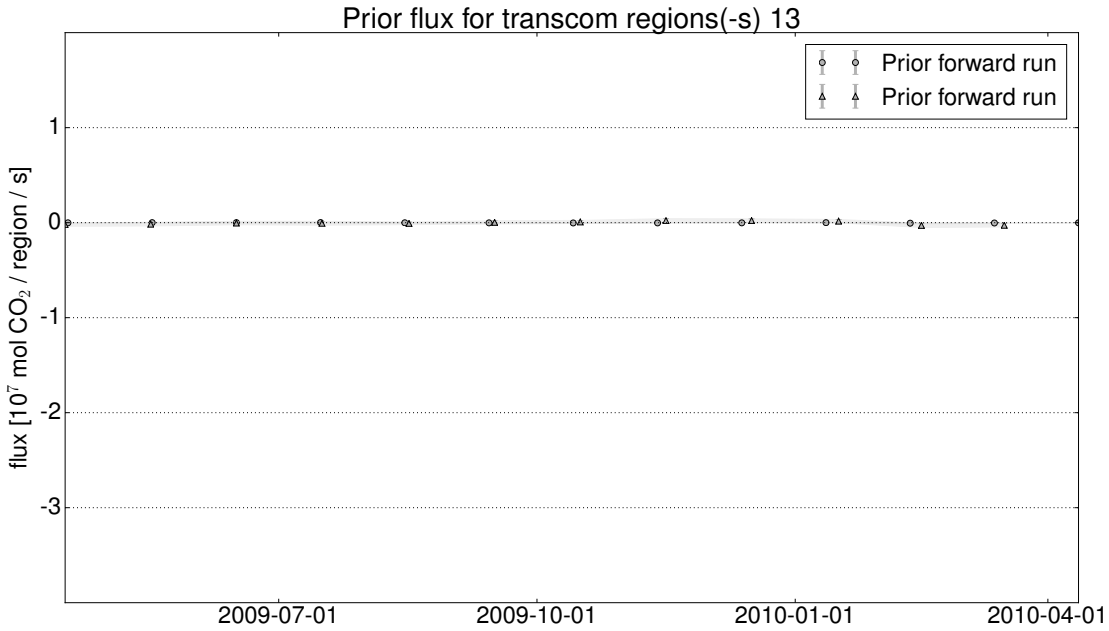


Figure A.13: Monthly prior flux and uncertainty for Tropical West Pacific.

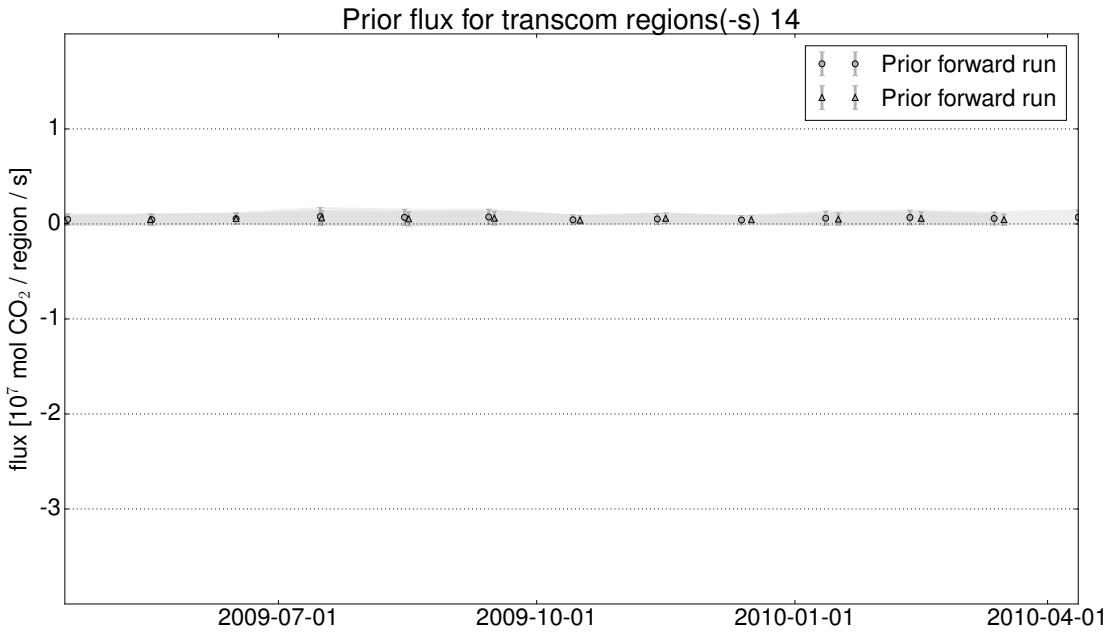


Figure A.14: Monthly prior flux and uncertainty for Tropical East Pacific.

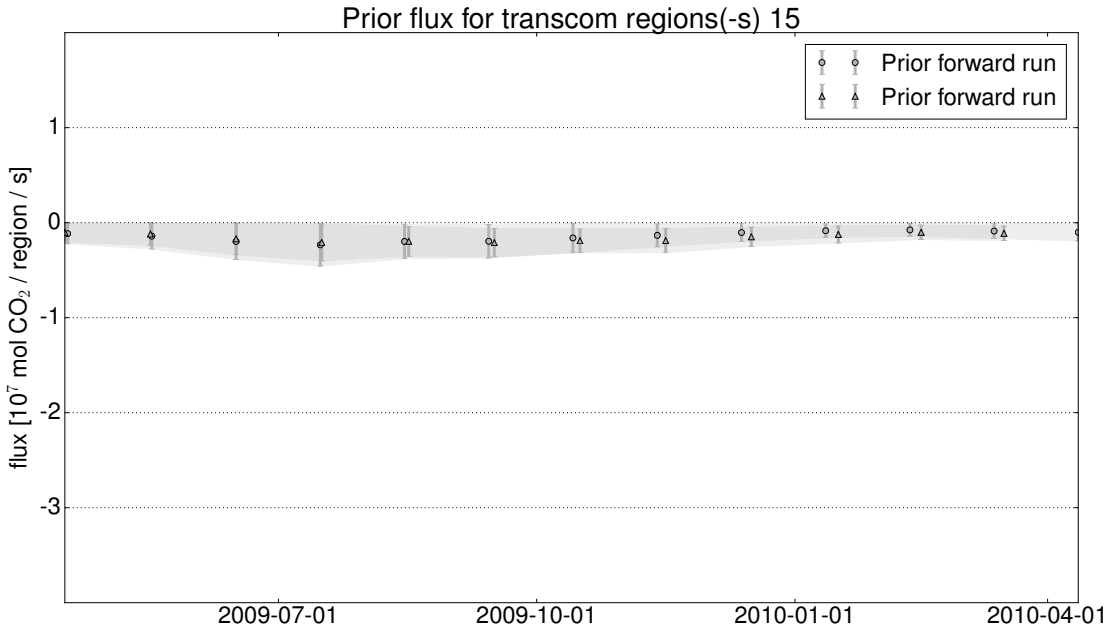


Figure A.15: Monthly prior flux and uncertainty for Temperate South Pacific.

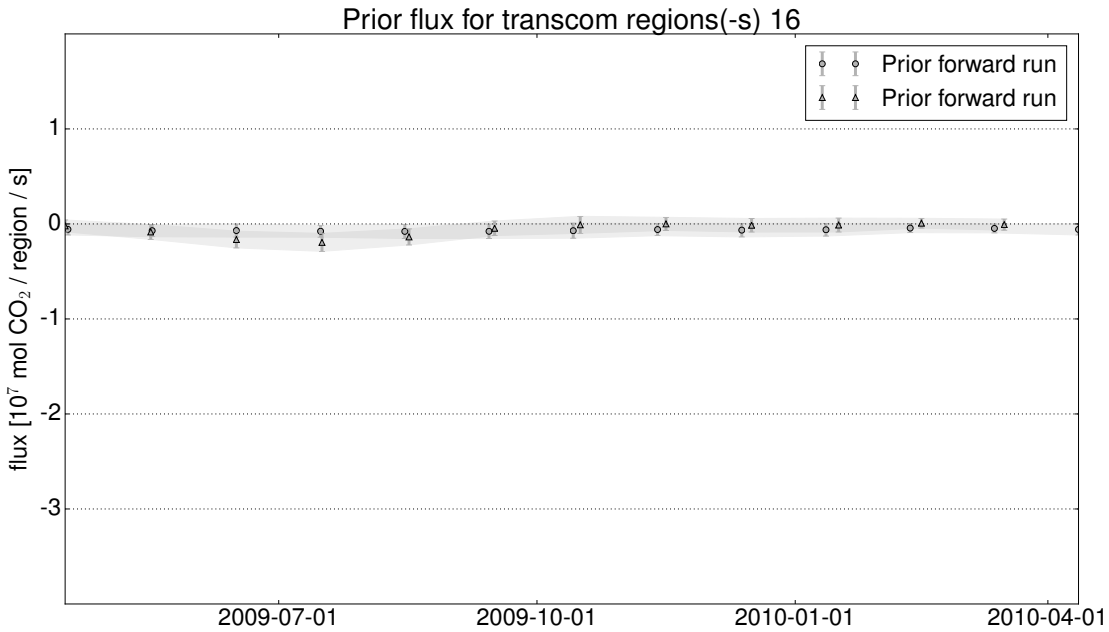


Figure A.16: Monthly prior flux and uncertainty for the Northern Ocean.

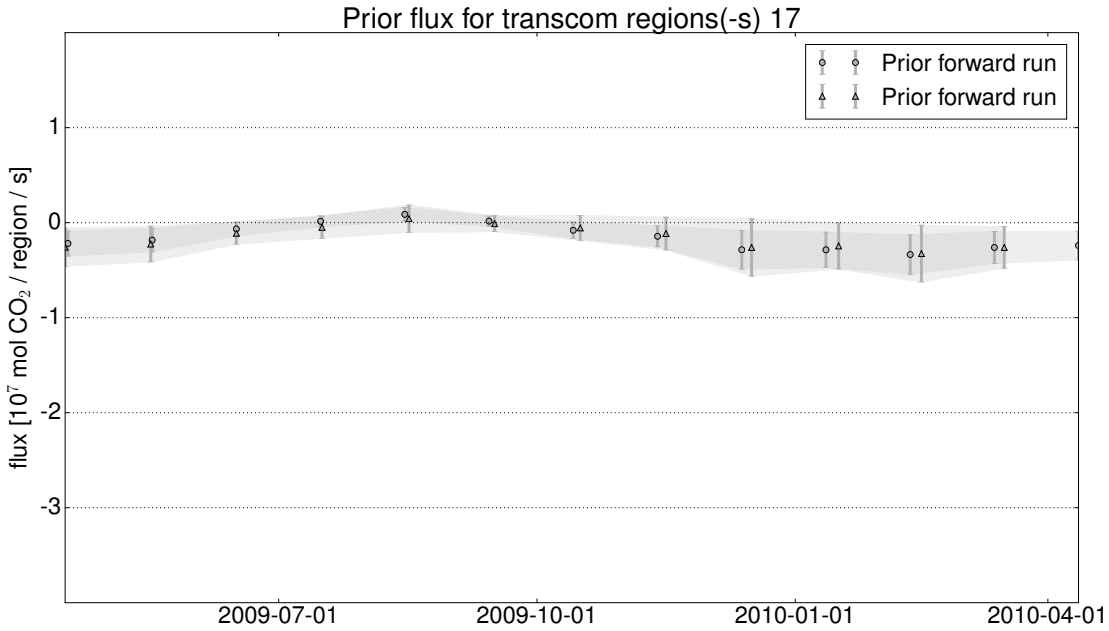


Figure A.17: Monthly prior flux and uncertainty for Temperate North Atlantic.

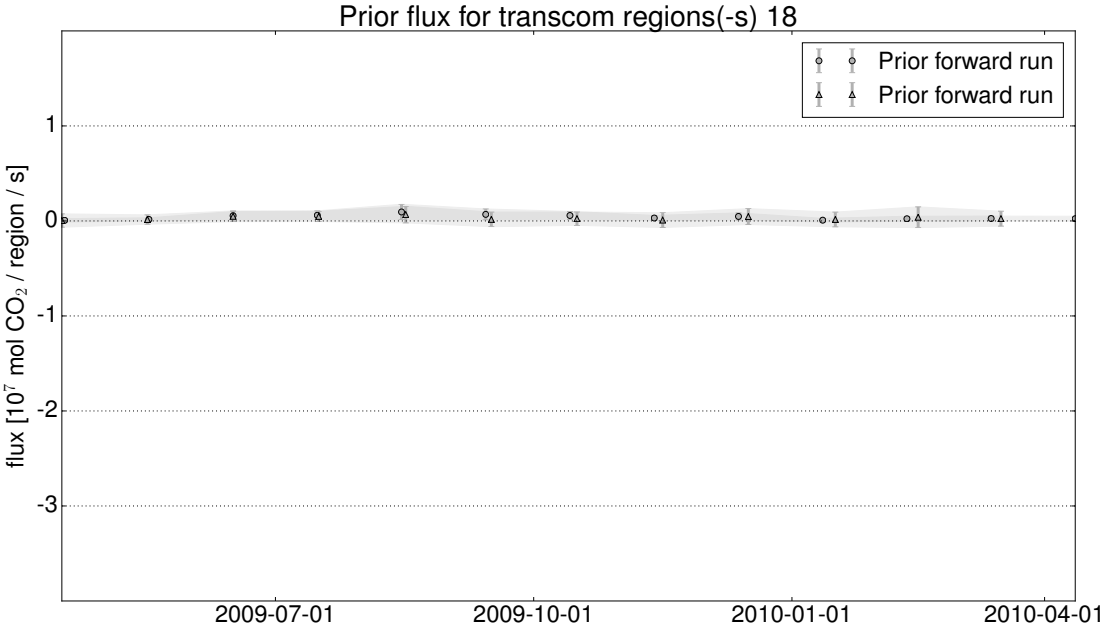


Figure A.18: Monthly prior flux and uncertainty for Tropical Atlantic.

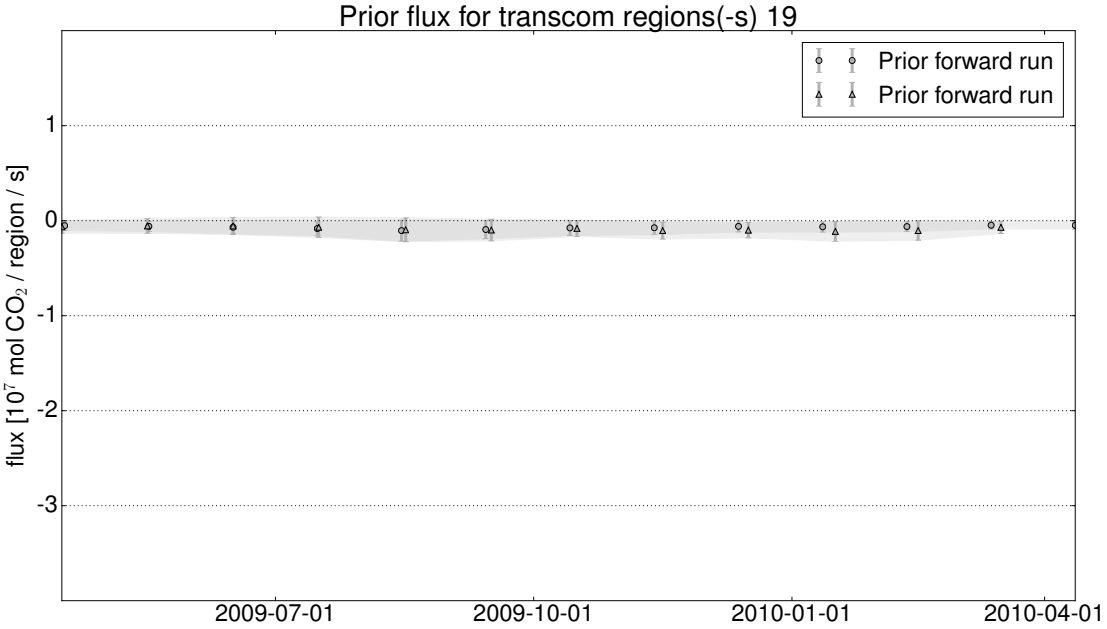


Figure A.19: Monthly prior flux and uncertainty for Temperate South Atlantic.

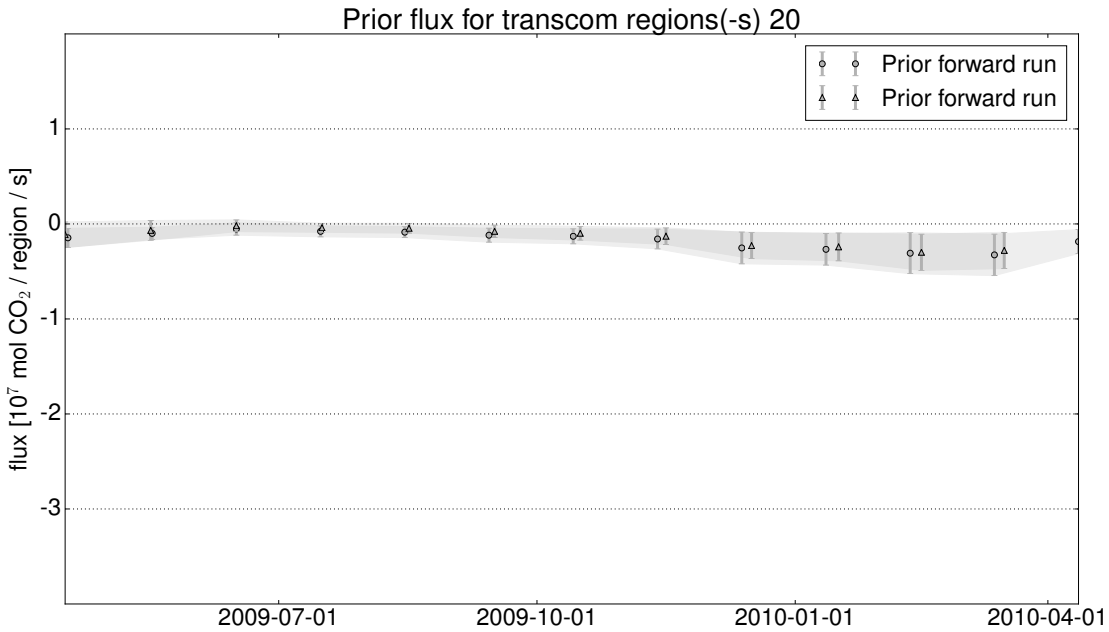


Figure A.20: Monthly prior flux and uncertainty for the Southern Ocean.

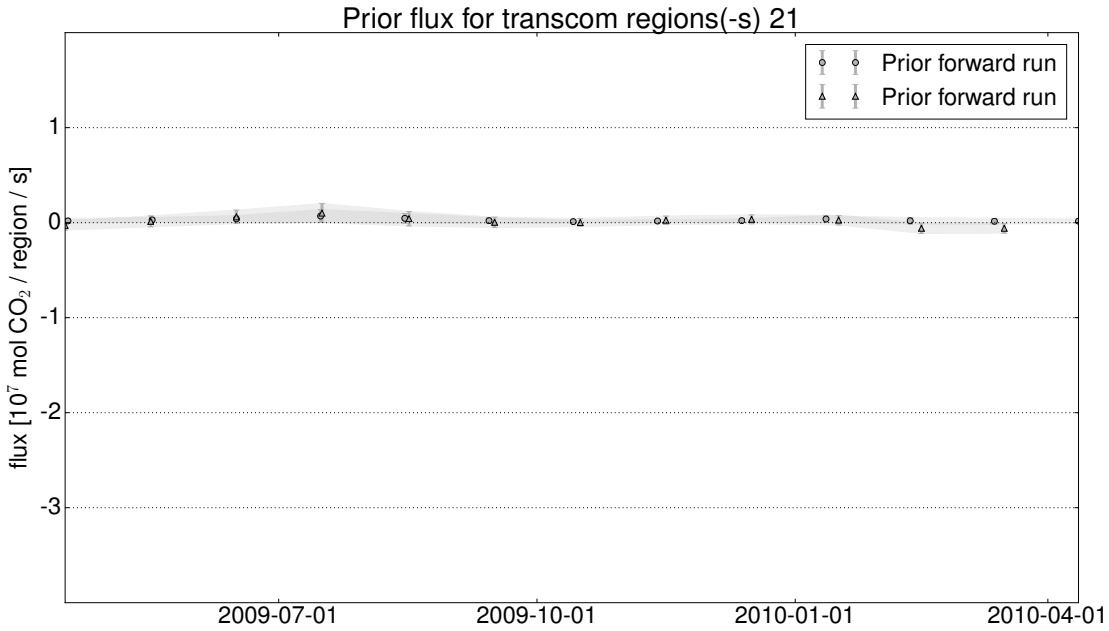


Figure A.21: Monthly prior flux and uncertainty for Tropical Indian Ocean.

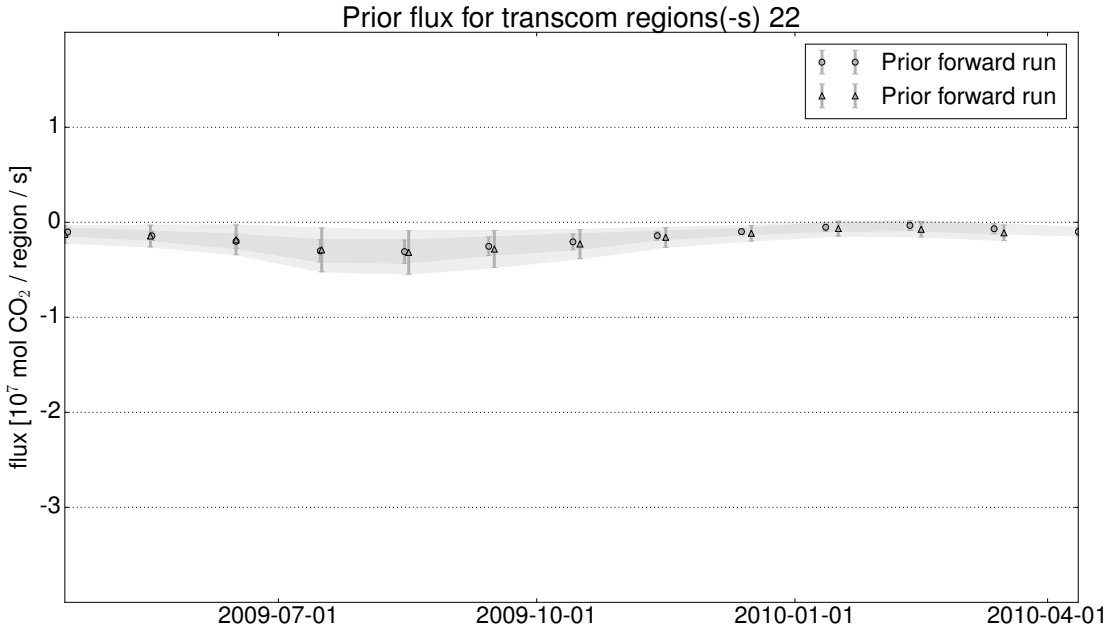


Figure A.22: Monthly prior flux and uncertainty for Temperate Indian Ocean.

Table A.1: Observation sites assimilated in the different runs or used in the resampling study, referenced by their NOAA sitecodes. Sites marked with * were removed as not representative of TM5-4DVar-input. Sites not assimilated in the base run are used to validate it. The Site column lists the sitecode, the Meas column gives the measurement type (f: discrete manual flask sampling, c: continuous in situ measurement, p: discrete automated flask sampling), the σ column gives the used representativeness error, the columns base, columns 2/cont, 1988, 2000 and 2010 list whether the site is assimilated in the given model run and the resampling column shows whether the site is used in the resampling study.

Site	Meas	σ [ppm]	base	2/cont	1988	2000	2010	resampling
ABP	f	0.75	yes	-	-	-	yes	-
ALT	c	2.50	yes	-	yes	yes	yes	-
ALT	f	1.50	yes	-	yes	yes	yes	-
AMT	c	3.00	yes	-	-	-	yes	-
AMT	f	3.00	-	-	-	-	-	-
AMT	p	3.00	yes	-	-	-	yes	-
ASC	f	0.75	yes	-	yes	yes	yes	-
ASK	f	1.50	yes	yes	-	yes	yes	-
AZR	f	1.50	yes	-	yes	yes	yes	-
BAL	f	7.50	yes	-	-	yes	yes	-
BAO	c	3.00	*	-	-	-	yes	yes
BAO	p	3.00	*	-	-	-	yes	yes
BGU	f	1.50	-	-	-	-	-	yes
BHD	f	0.75	yes	-	-	yes	yes	-
BKT	f	7.50	yes	-	-	-	yes	-
BME	f	1.50	yes	-	-	yes	yes	-
BMW	f	1.50	yes	-	-	yes	yes	-
BNE	p	2.00	-	-	-	-	-	yes
BRA	c	3.00	-	-	-	-	yes	yes
BRW	f	1.50	-	-	yes	yes	yes	yes
BRW	c	2.50	-	-	yes	yes	yes	yes
BSC	f	7.50	yes	-	-	yes	yes	-
CAR	p	2.00	-	-	-	-	-	yes
CBA	f	1.50	yes	-	yes	yes	yes	-
CDL	c	3.00	-	-	-	-	yes	yes
CFA	f	2.50	yes	yes	-	yes	yes	-
CGO	f	0.75	-	-	yes	yes	yes	yes
CGO	c	2.50	-	-	yes	yes	yes	yes
CHM	c	3.00	-	-	-	-	yes	yes
CHR	f	0.75	yes	-	yes	yes	yes	-
CIB	f	1.50	-	-	-	-	-	yes
CMA	p	2.00	-	-	-	-	-	yes
CPT	f	2.50	-	-	-	-	-	-
CPT	c	2.50	-	-	-	yes	yes	yes
CRI	f	0.75	-	-	-	-	-	yes
CRZ	f	0.75	yes	-	-	-	yes	-
CYA	f	0.75	yes	-	-	yes	yes	-
DND	p	2.00	-	-	-	-	-	yes

Continued on next page

Table A.1 – continued from previous page

Site	Meas	σ [ppm]	base	2/cont	1988	2000	2010	resampling
DRP	f	0.75	-	-	-	-	-	yes
EGB	c	3.00	yes	-	-	-	yes	yes
EIC	f	7.50	-	-	-	yes	yes	-
ESP	c	3.00	-	-	-	yes	yes	yes
ESP	p	3.00	-	-	-	-	-	yes
ESP	f	3.00	-	-	-	-	-	-
EST	c	1.50	-	-	-	-	-	yes
ETL	c	3.00	yes	yes	-	-	yes	yes
ETL	p	3.00	-	-	-	-	-	yes
FIK	f	1.50	-	-	-	-	-	yes
FNS	c	1.50	-	-	-	-	-	yes
FSD	c	1.50	-	-	-	-	-	yes
GMI	f	1.50	-	-	yes	yes	yes	yes
HBA	f	0.75	yes	yes	yes	yes	yes	-
HDP	c	3.00	-	-	-	-	yes	yes
HEI	c	1.50	-	-	-	-	-	yes
HIL	p	2.00	-	-	-	-	-	yes
HPB	f	2.50	yes	-	-	-	yes	-
HUN	f	7.50	-	-	-	yes	yes	yes
HUN	c	7.50	-	-	-	-	-	-
ICE	f	1.50	yes	-	-	yes	yes	-
IZO	f	1.50	yes	-	-	yes	yes	-
IZO	c	1.50	-	-	-	-	-	-
JFJ	c	1.50	-	-	-	-	-	yes
KEY	f	2.50	yes	-	yes	yes	yes	-
KUM	f	1.50	yes	-	yes	yes	yes	-
KZD	f	2.50	yes	-	-	yes	yes	-
KZM	f	2.50	yes	-	-	yes	-	-
LEF	c	3.00	-	-	-	-	yes	yes
LEF	f	2.50	-	-	-	-	-	yes
LEF	p	3.00	-	-	-	-	yes[1]	yes
LJO	f	1.50	-	-	-	-	-	yes
LLB	c	3.00	yes	-	-	-	yes	-
LLB	f	3.00	-	-	-	-	-	-
LLN	f	2.50	yes	-	-	-	yes	-
LMP	f	2.50	yes	-	-	-	yes	-
LPO	f	1.50	-	-	-	-	-	yes
LUT	c	1.50	-	-	-	-	-	yes
MAA	f	0.75	yes	-	-	yes	yes	-
MEX	f	2.50	-	-	-	-	-	yes
MHD	f	2.50	yes	yes	-	yes	yes	-
MID	f	1.50	yes	-	yes	yes	yes	-
MKN	f	2.50	yes	-	-	-	yes	-
MLO	f	1.50	yes	yes	yes	yes	yes	-
MLO	c	0.75	-	–	yes	yes	yes	-

Continued on next page

[1] only afternoon samples

Table A.1 – continued from previous page

Site	Meas	σ [ppm]	base	2/cont	1988	2000	2010	resampling
MQA	f	0.75	yes	-	-	yes	yes	-
NHA	p	2.00	-	-	-	-	-	yes
NMB	f	2.50	yes	yes	-	yes	yes	-
NWR	f	1.50	-	-	yes	yes	yes	yes
NWR	p	1.50	-	-	yes	yes	yes	yes
NWR	c	1.50	-	-	yes	yes	yes	yes
OBN	f	7.50	yes	yes	-	-	-	-
ORL	p	1.50	-	-	-	-	-	-
ORL	f	1.50	-	-	-	-	-	-
OTA	f	2.50	-	-	-	-	-	yes
OXK	f	2.50	yes	-	-	-	yes	-
PAL	f	2.50	yes	-	-	yes	yes	-
PAL	c	2.50	-	-	-	-	-	-
PDM	f	1.50	-	-	-	-	-	yes
PFA	p	2.00	-	-	-	-	-	yes
POC	f	0.75	yes	-	yes	yes	yes	-
PSA	f	0.75	yes	-	yes	yes	yes	-
PTA	f	7.50	yes	-	-	yes	yes	-
RBA	c	1.50	-	-	-	-	-	yes
RPB	f	1.50	yes	yes	yes	yes	yes	-
RTA	p	2.00	-	-	-	-	-	yes
SBL	c	3.00	yes	-	-	-	-	-
SAN	f	2.00	-	-	-	-	-	yes
SAN	p	2.00	-	-	-	-	-	yes
SCA	p	2.00	-	-	-	-	-	yes
SCT	c	3.00	*	-	-	-	yes	yes
SCT	p	3.00	*	-	-	-	-	yes
SEY	f	0.75	yes	yes	yes	yes	yes	-
SGP	f	2.50	yes	-	-	-	yes	-
SGP	c	3.00	yes	-	-	-	yes	-
SGP	p	2.50	-	-	-	-	-	-
SHM	f	2.50	yes	-	yes	yes	yes	-
SMO	f	1.50	yes	yes	yes	yes	yes	-
SMO	c	0.75	yes	yes	yes	yes	yes	-
SNP	c	3.00	-	-	-	-	yes	yes
SPL	c	3.00	yes	-	-	-	yes	-
SPO	f	1.50	yes	-	yes	yes	yes	-
SPO	c	0.75	yes	-	yes	yes	yes	-
STM	f	1.50	yes	-	yes	yes	yes	-
STR	p	3.00	yes	-	-	-	yes	-
SUM	f	1.50	yes	-	-	yes	yes	-
SYO	f	0.75	-	-	yes	yes	yes	yes
SYO	c	0.75	-	-	-	-	-	yes
TAP	f	7.50	yes	-	-	yes	yes	-
TDF	f	0.75	yes	-	-	yes	yes	-
TGC	p	2.00	-	-	-	-	-	yes

Continued on next page

Table A.1 – continued from previous page

Site	Meas	σ [ppm]	base	2/cont	1988	2000	2010	resampling
THD	f	2.50	yes	-	-	-	yes	-
THD	p	2.50	-	-	-	-	-	-
TOT	c	1.50	-	-	-	-	-	yes
TRN	c	1.50	-	-	-	-	-	yes
UTA	f	2.50	yes	-	-	yes	yes	-
UUM	f	2.50	yes	-	-	yes	yes	-
WBI	c	3.00	*	yes	-	-	yes	yes
WBI	p	3.00	*	yes	-	-	yes	yes
WGC	c	3.00	*	-	-	-	yes	yes
WGC	p	3.00	*	-	-	-	yes	yes
WIS	f	2.50	yes	-	-	yes	yes	-
WKT	c	3.00	*	-	-	-	yes	yes
WKT	p	3.00	*	-	-	-	yes	yes
WKT	f	2.50	*	-	-	-	-	yes
WLG	f	1.50	yes	yes	-	yes	yes	-
WPC	f	1.50	-	-	-	-	-	-
WSA	c	3.00	yes	yes	-	-	yes	-
ZEP	f	1.50	yes	yes	-	yes	yes	-

A.3 Learjet flight information

Information about the learjet data used in this study to assess the viability of assimilating column data from the TCCON network.

```
CO2 data from the Learjet flight over the SGP ARM Site on July 31, 2009.
aircraft_info:                  NASA Glenn Lear-25 (5-13 km altitude) and Cessna 210 (0.3-5 km altitud
flight_date:                    2009-07-31
aircraft_floor_time_UTC:        2009-07-31 16:01:00
aircraft_start_time_UTC:        2009-07-31 14:37:00
aircraft_stop_time_UTC:         2009-07-31 17:31:00
altitude_source:                Aircraft
pressure_source:                Radiosonde SGPCCG
temperature_source:             Radiosonde SGPCCG
h2o_profile_source:             Radiosonde SGPCCG
co2_profile_source:             NOAA Flask Samplers
aircraft_co2_error_2sigma_ppm: 0.31
aircraft_ceiling_m:             11233.4
aircraft_floor_m:               457.2
co2_surface_source:             SGP Flasks
h2o_surface_source:             FTS RH Measurement
TCCON_site_name:                Lamont
TCCON_site_longitude_E:         -97.486
TCCON_site_latitude_N:          36.604

CO2 data from the Learjet flight over the SGP ARM Site on August 02, 2009.
```

```
aircraft_info:                  NASA Glenn Lear-25 (5-13 km altitude) and Cessna 210 (0.3-5 km
flight_date:                    2009-08-02
aircraft_floor_time_UTC:        2009-08-02 16:30:00
aircraft_start_time_UTC:        2009-08-02 15:05:00
aircraft_stop_time_UTC:         2009-08-02 17:57:00
altitude_source:                Aircraft
pressure_source:                Radiosonde SGPCCG
temperature_source:             Radiosonde SGPCCG
h2o_profile_source:             Radiosonde SGPCCG
co2_profile_source:             NOAA Flask Samplers
aircraft_co2_error_2sigma_ppm: 0.31
aircraft_ceiling_m:             12831.8
aircraft_floor_m:               457.2
co2_surface_source:             SGP Flasks
h2o_surface_source:             FTS RH Measurement
TCCON_site_name:                Lamont
TCCON_site_longitude_E:         -97.486
TCCON_site_latitude_N:          36.604

CO2 data from the Learjet flight over the SGP ARM Site on August 03, 2009.
aircraft_info:                  NASA Glenn Lear-25 (5-13 km altitude) and Cessna 210 (0.3-5 km
flight_date:                    2009-08-03
aircraft_floor_time_UTC:        2009-08-03 16:40:00
aircraft_start_time_UTC:        2009-08-03 15:14:00
aircraft_stop_time_UTC:         2009-08-03 18:00:00
altitude_source:                Aircraft
pressure_source:                Radiosonde SGPCCG
temperature_source:             Radiosonde SGPCCG
h2o_profile_source:             Radiosonde SGPCCG
co2_profile_source:             NOAA Flask Samplers
aircraft_co2_error_2sigma_ppm: 0.31
aircraft_ceiling_m:             12855.5
aircraft_floor_m:               457.2
co2_surface_source:             SGP Flasks
h2o_surface_source:             FTS RH Measurement
TCCON_site_name:                Lamont
TCCON_site_longitude_E:         -97.486
TCCON_site_latitude_N:          36.604

CO2 data from the Learjet flight over the SGP ARM Site on July 18, 2010.
aircraft_info:                  Learjet
flight_date:                    2010-07-18
aircraft_floor_time_UTC:        2010-07-18 18:25:14
aircraft_start_time_UTC:        2010-07-18 16:15:39
aircraft_stop_time_UTC:         2010-07-18 20:27:54
altitude_source:                Learjet
pressure_source:                Radiosonde SGPCCP
temperature_source:             Radiosonde SGPCCP
h2o_profile_source:             Radiosonde SGPCCP
```

```
co2_profile_source:             Stephanie Vay's CO2
aircraft_co2_error_2sigma_ppm: 0.3
aircraft_ceiling_m:             12932.0
aircraft_floor_m:               410.0
co2_surface_source:             Extrapolated from aircraft measurements
h2o_surface_source:             FTS RH Measurement
TCCON_site_name:                Lamont
TCCON_site_longitude_E:         -97.486
TCCON_site_latitude_N:          36.604
```

A.4 The EnSRF Toy Model

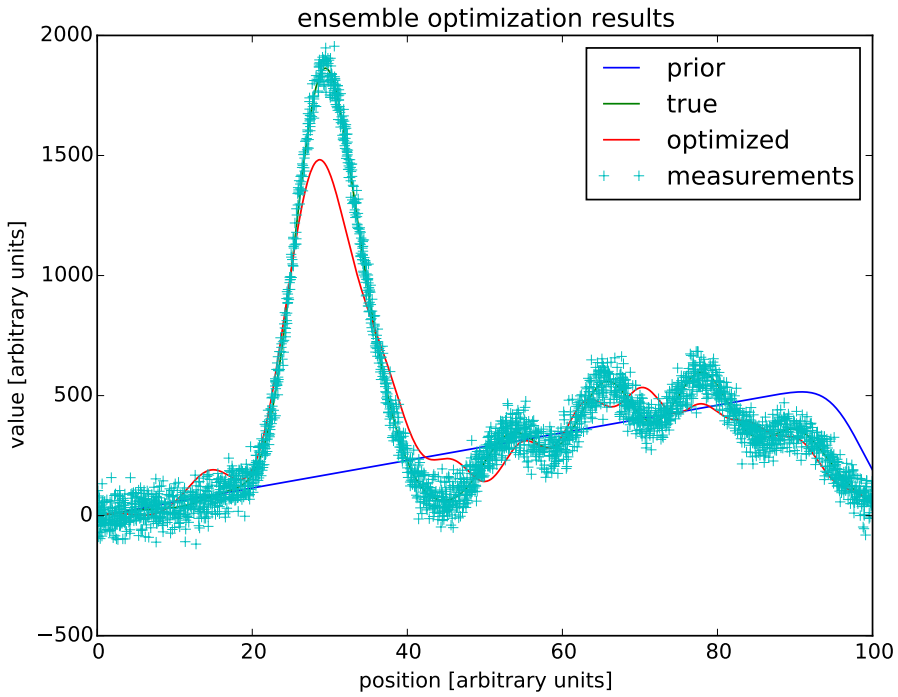


Figure A.23: Output from the EnSRF toy model in appendix A.4 using a linear model. The example shows how EnSRF optimization deviates from the true value even with a dense measurement coverage when the model parameters have very different weight. It motivates the need to precondition the state.

This section contains a toy implementation of an Ensemble Square Root Filter (EnSRF) written in Wisp, an indentation based syntax for Guile Scheme. It is used in section 2.5.3 as motivation for the need to precondition the state vector.

```
# -*- wisp -*-
# save this file in the wisp folder and execute it in place.
exec guile -L . --language=wisp -e '(@@ (ensrf) main)' -s "$0" "$@"
; !#

;; Simple Ensemble Square Root Filter to estimate function parameters
;; based on measurements.

;; Provide first guess parameters x^b and measurements y^0 to get
;; optimized parameters x^a.

;; This EnSRF toy model uses Wisp syntax, also known as SRFI-119. See
;; http://draketo.de/light/english/wisp-lisp-indentation-preprocessor
;; http://srfi.schemers.org/srfi-119/srfi-119.html
```

Listing 1: Basic file execution and information.

```
;; Method
;; x^b = '(...) ; first guess of the parameters
;; P = '((...) (...) ...) ; parameter covariance
;; y^0 = '(...) ; observations
;; R = '((...) (...) ...) ; observation covariance
;; H: H(x) → y ; provide modelled observations for the given parameters.
;; with N ensemble members (i=1, ... N) drawn from the state x^b:
;; For each measurement y^0_j:
;;     x'^b: X = 1/√(N-1)(x'b_1, ..., x'b_N)^T
;;     with P = XX^T ; in the simplest case x'^b are gaussian
;;                     distributed with standard distribution from
;;                     square root of the diagonals.
;;     x_i = x^b + x'^b_i
;;     H(x^b_i) = H(x^b + x'^b_i)
;;     H(x^b) = (1/N)·Σ H(x^b + x'^b_i)
;;     H(x'^b_i) = H(x^b + x'_i) - H(x^b)
;;     HPHt = 1/(N-1)(H(x'_1), ..., H(x'_N))(H(x'1), ..., H(x'N))T
;;     PHt = 1/(N-1)(x'_1, ..., x'_N)(H(x'1), ..., H(x'N))T
;;     K = PHt*(HPHt + R)^-1
;;     x^a = x^b + K(y^0_j - H(x^b))
;;     α = (1 + √(R/(HPHt+R)))^-1
;;     x'^a = x'^b - αK·H(x'^b)
```

Listing 2: The EnSRF algorithm.

```
define-module : ensrf
use-modules : srfi srfi-42 ; list-ec
use-modules
  : ice-9 popen
        . #:select : open-output-pipe close-pipe

; seed the random number generator
set! *random-state* : random-state-from-platform
```

Listing 3: Module definition, imports and preparation

```
define : make-diagonal-matrix-with-trace trace
         let : : dim : length trace
             list-ec (: i dim)
               list-ec (: j dim)
                 if : = i j
                      list-ref trace i
                      . 0.0

define : make-covariance-matrix-from-standard-deviations stds
         make-diagonal-matrix-with-trace : map (lambda (x) (expt x 2)) stds

define : mean l
       . "Calculate the average value of l (numbers)."
       / : apply + l
           length l


define : standard-deviation l
       . "Calculate the standard deviation of list l (numbers)."
       let : : l_mean : mean l
         sqrt
           / : sum-ec (: i l) : expt {i - l_mean} 2
             . {(length l) - 1}

define : standard-deviation-from-deviations . l
       . "Calculate the standard deviation from a list of deviations (x - x_mean)."
       sqrt
         / : sum-ec (: i l) : expt i 2
           . {(length l) - 1}

define* : write-multiple . x
        . "Helper to avoid suffering from write-newline-typing."
        map : lambda (x) (write x) (newline)
            . x
```

Listing 4: Helper functions.

```
;; Start with the simple case: One variable and independent observations (R diagonal)
;; First define a truth
define x^seed '(0.5 0.6 7 0.1 0.7 0.9 0.8 0.4)
define x^true
  append-ec (: i (length x^seed))
    list-ec (: j x^seed) : * j : list-ref x^seed i
;; And add an initial guess of the parameters
define x^b : append-ec (: i (length x^seed))  '(1 1 1 1 1 1 1 1)
define P
  make-covariance-matrix-from-standard-deviations
    append-ec (: i (length x^seed)) '(0.5 0.1 0.3 0.1 0.2 0.2 0.2 0.2)

;; Then generate observations
define y^0-num 3000
define y^0-pos-max 100
;; At the positions where they are measured. Drawn randomly to avoid
;; giving an undue weight to later values.
define y^0-pos : list-ec (: i y^0-num) : * (random:uniform) y^0-pos-max
```

Listing 5: Setup of model and observation locations.

```
;; We need an observation operator to generate observations from true values
define : H x pos
       . "Observation operator. It generates modelled observations from the input.

x are parameters to be optimized, pos is another input which is
not optimized. For plain functions it could be the position of
the measurement on the x-axis. We currently assume absolute
knowledge about the position.
"
       let*
           : len : length x
             ystretch y^0-pos-max
             x-pos : list-ec (: i len) : * ystretch {{i + 0.5} / {len + 1}}
           apply +
                list-ec (: i len)
                     * : list-ref x i
                         . pos
                         exp
                           -
                             expt
                               / {pos - (list-ref x-pos i)} {ystretch / 20}
                               . 2
```

Listing 6: The observation operator.

```
;; We start with true observations which we will disturb later to get
;; the equivalent of measured observations
define y^true : list-ec (: i y^0-pos) : H x^true i
;; now we disturb the observations with a fixed standard deviation.
;; This assumes uncorrelated observations.
define y^0-std 50
define y^0 : list-ec (: i y^true) : + i : * y^0-std : random:normal
;; and define the covariance matrix. This assumes uncorrelated observations.
define R : make-covariance-matrix-from-standard-deviations : list-ec (: i y^0-num) y^0-std

;; Alternative: define observations
;; define y^0-mean 0.8
;; The actual observations
;; define y^0 : list-ec (: i y^0-num) : + y^0-mean : * y^0-std : random:normal
```

Listing 7: Generate observations and observation uncertainty.

```
observations-to-process y
observation-variances : list-ec (: i (length y)) : list-ref (list-ref R i) i
observation-positions y-pos
x^b x
x-deviations
    list-ec (: i N)
      list-ec (: j (length x))
          * : random:normal
              sqrt : list-ref (list-ref P j) j ; only for diagonal P!
```

Listing 8: EnSRF Step, part 1: Start of one EnSRF step.

```
y_cur : car observations-to-process
R_cur : car observation-variances
y-pos_cur : car observation-positions
Hx^b_i
    list-ec (: i x-deviations)
        H
          list-ec (: j (length i))
              + (list-ref x^b j) (list-ref i j)
          . y-pos_cur
Hx^b
   / : sum-ec (: i Hx^b_i) i
     . N
Hx^b-prime
    list-ec (: i N)
        - : list-ref Hx^b_i i
          . Hx^b
```

Listing 9: EnSRF Step, part 2: Processing one observation.

```
HPHt
   / : sum-ec (: i Hx^b-prime) {i * i}
     . {N - 1}
PHt
   list-ec (: j (length x^b))
      ; for each x^b_i multiply the state-element
      ; and model-deviation for all ensemble members.
      * {1 / {N - 1}}
          sum-ec (: i N)
            * : list-ref (list-ref x-deviations i) j
                list-ref Hx^b-prime i
K : list-ec (: i PHt) {i / {HPHt + R_cur}}
```

Listing 10: EnSRF Step, part 3: Calculating the Kalman gain matrix.

```
x^a
  list-ec (: j (length x^b))
    + : list-ref x^b j
      * : list-ref K j
        . {y_cur - Hx^b}
α-weight-sqrt : sqrt {R_cur / {HPHt + R_cur}}
α {1 / {1 + α-weight-sqrt}}
x^a-deviations
  list-ec (: i N) ; for each ensemble member
    list-ec (: j (length x^b)) ; and each state variable
      - : list-ref (list-ref x-deviations i) j
        * α
          list-ref K j
          list-ref Hx^b-prime i
```

Listing 11: EnSRF Step, part 4: Update state and deviations.

```
define : EnSRF H x P y R y-pos N
       . "Observation function H, parameters x,
parameter-covariance P, observations y, observation covariance R
and number of ensemble members N.

Limitations: y is a single value. R and P are diagonal.
"
       let step
         :
           <<wisp-ensrt-start>>
         cond
            : null? observations-to-process
              list x^b x-deviations
            else
               let*
                 :
                   <<wisp-ensrt-oneobs>>
                   <<wisp-ensrt-gain>>
                   <<wisp-ensrt-update>>
                 step
                   cdr observations-to-process
                   cdr observation-variances
                   cdr observation-positions
                   . x^a
                   . x^a-deviations
```

Listing 12: The EnSRF implementation.

```
format #t "x⁰: ~A ± ~A\nx:  ~A ± ~A\nx^t:~A\ny:  ~A ± \ny⁰: ~A ± ~A\nnoise: ~A\n"
           . x^b
           list-ec (: i (length x^b)) : list-ref (list-ref P i) i
           . x-opt
           list-ec (: i (length x-opt))
              apply standard-deviation-from-deviations
                list-ec (: j x-deviations) : list-ref j i
           . x^true
           * {1 / (length y⁰)} : apply + : map (lambda (x) (H x-opt x)) y⁰-pos
           mean y⁰
           standard-deviation y⁰
           . y⁰-std
```

Listing 13: Show EnSRF information.

```
; plot the result
let : : port : open-output-pipe "python"
  format port "import pylab as pl\n"
  format port "y0 = [float(i) for i in '~A'[1:-1].split(' ')]\n" y^0
  format port "ypos = [float(i) for i in '~A'[1:-1].split(' ')]\n" y^0-pos
  format port "yinit = [float(i) for i in '~A'[1:-1].split(' ')]\n" : list-ec (: i y^0-pos) : H x^b i
  format port "ytrue = [float(i) for i in '~A'[1:-1].split(' ')]\n" : list-ec (: i y^0-pos) : H x^true i
  format port "yopt = [float(i) for i in '~A'[1:-1].split(' ')]\n" : list-ec (: i y^0-pos) : H x-opt i
  format port "pl.plot(*zip(*sorted(zip(ypos, yinit))), label='prior')\n"
  format port "pl.plot(*zip(*sorted(zip(ypos, ytrue))), label='true')\n"
  format port "pl.plot(*zip(*sorted(zip(ypos, yopt))), label='optimized')\n"
  format port "pl.plot(*zip(*sorted(zip(ypos, y0))), marker='+', linewidth=0, label='measurements')\n"
  format port "pl.legend()\n"
  format port "pl.xlabel('position [arbitrary units]')\n"
  format port "pl.ylabel('value [arbitrary units]')\n"
  format port "pl.title('ensemble optimization results')\n"
  format port "pl.show()\n"
  format port "exit()\n"
  close-pipe port
```

Listing 14: Plot EnSRF results.

```
<<wisp-header>>
<<wisp-math>>
<<wisp-setup>>
<<wisp-helpers>>
<<wisp-input>>
<<wisp-obsoperator>>
<<wisp-observations>>
<<wisp-ensrt>>

define : main args
    let*
      : optimized : EnSRF H x^b P y^0 R y^0-pos 30
        x-opt : list-ref optimized 0
        x-deviations : list-ref optimized 1
      <<wisp-show>>
      <<wisp-plot>>
```

Listing 15: Collect the parts into the EnSRF toy model.