Regional carbon balance: flux optimization and network design

Dissertation

zur Erlangung des akademischen Grades doctor rerum naturalium

(Dr. rer. nat.)

vorgelegt dem Rat der Chemisch-Geowissenchaftlichen Fakultät

der

Friedrich-Schiller-Universität Jena

von Dipl. Phys. Kountouris, Panagiotis

geboren am 29. Januar 1982

in Athen, Griechenland

Gutachter:

1. ……

2. ……

Tag der öffentlichen Verteidigung: .....

Abstract

Long term observations of atmospheric greenhouse gas measuring stations, located at representative regions over the continent, improve our understanding of greenhouse gas sources and sinks. These dry mole fraction measurements can be linked to surface fluxes by atmospheric transport inversions. A number of stations is already deployed within the European domain, and, in the framework of ICOS (Integrated Carbon Observation System, an EU infrastructure project), more observing stations are to be deployed. The design of these observational networks amounts to the optimization of the atmospheric constraint on specific targeted quantities. A quantitative network design study is required to perform this optimization and to assess potential observing networks.

For this reason a regional inverse modeling framework was set up that derives biosphere-atmosphere exchange fluxes at regional scales using CO_2 measurements from tall towers, ground stations, and mountain sites. The modeling framework consists of the following components: the global transport model TM3, the regional Stochastic Time-Inverted Lagrangian Transport model (STILT; a Lagrangian particle dispersion model), the Vegetation Photosynthesis and Respiration Model (VPRM; a diagnostic biosphere model of fluxes from the land biosphere), gridded emissions from fossil fuel burning, ocean-atmosphere exchange fluxes, and a Bayesian inversion scheme.

The inverse problem is an ill-posed problem since a limited number of observations is used to determine a much larger number of unknowns (e.g. spatially and temporally resolved fluxes). The Bayesian approach offers a rigorous framework to solve this problem by making use of a-priori flux estimates. Moreover, this framework accounts also for errors from both, the a-priori information and the observations, and derives a posterior estimate which is a balance between these two constraints. It is essential to thoroughly quantify these uncertainties as they determine the relative weight of the constraints.

Atmospheric CO_2 measurements from tall towers are strongly influenced by near field surface fluxes (areas within a radius up to approximately 100 km around the station location). Using a high resolution system is imperative to resolve the flux variability at small scales that global systems can not capture. Hence high spatial resolution was used to couple the atmospheric transport model with biogenic fluxes (i.e. 0.25°). The aim of this thesis is to quantitatively assess the impact of different network configurations and for that an uncertainty reduction analysis is being carried out. Such an analysis is strongly dependent on the assumed prior error structure. Realistic uncertainties, which will allow for an objective network assessment, should be assumed.

Quantitative treatment of uncertainties is paramount to quantitatively estimating biosphere-atmosphere fluxes. To better describe flux uncertainties in the inversion system this thesis first studies the flux error structure, and explains how these uncertainties are distributed spatially and temporally. Fluxes from three biosphere models were used and compared against flux observations from 53 Eddy covariance flux towers and from an aircraft campaign. Spatial and temporal autocorrelations of the daily model-data flux residuals can be approximated by an exponentially decay error model. The autocorrelation of daily model-data flux residuals showed e-folding temporal correlations of 30 days. Spatial autocorrelation e-folding lengths were found to be only up to 40 km, whereas model-model residual spatial autocorrelation lengths were found

to be up to some hundreds of km. Propagating the model-data error structure and up-scaling to European domain results in an annual domain-wide uncertainty of 0.15 $GtCy^{-1}$, which is a factor of 2 smaller than uncertainties derived from an ensemble of modeled flux retrievals. Therefore, this thesis suggests an appropriate error inflation using two uncertainty terms. The first term reflects the spatiotemporal error structure and the second accounts for a potential large-scale bias. The bias term has no temporal variability and spatially follows a predefined flux field (i.e. respiration or NEE fluxes).

This quantitative and data driven information on the flux error structure is implemented in the inversion system and tested intensively. A synthetic experiment was performed using two different biosphere models, one to produce the a-priori flux field, and the other to provide fluxes that served as a "known truth". This experiment allows to quantitatively assess the system's ability to correct fluxes at different spatial and temporal scales. A very good agreement was found between retrieved and known fluxes at spatially aggregated regions down to country level and for annual and monthly temporal scales. Subsequently, real data atmospheric inversions were carried out using different flux error structures and sensitivity tests. The estimated carbon sink, ranges between 0.23 and 0.38 GtC y^{-1}. Validation of the monthly aggregated modeled fluxes at grid scale, against eddy covariance flux measurements, showed a very good agreement with a 52% bias reduction and 24% reduction in the standard deviation of the residuals.

A network design study is conducted after having tested the performance of the inversion system. Different network configurations were realized using the ICOS current and future stations and two inversion systems, from Jena (TM3-STILT) and from the Netherlands (VUA) provided uncertainty reduction estimates on the terrestrial fluxes. The spatial distribution of the uncertainty reduction significantly differs among the inversion systems showing a more localized effect for the TM3-STILT system. Annual uncertainty reduction scores at country scale differ up to 20% for the different inversion systems, but show a similar pattern across countries. The achieved annual domain-averaged uncertainty reduction for the current ICOS network is estimated to be between 10.2% and 12.5% according to the different systems. Overall the ICOS network expansion is expected to reduce the average domain-wide flux uncertainties within a range of 11.8% and 17.4% (for VUA and TM3-STILT systems, respectively).

Atmospheric observations exhibit measurement uncertainties, including potential bias errors. For CO_2, the WMO recommends keeping those bias errors below 0.1 ppm. The inversion system was used to assess to what extend atmospheric inverse estimation of sources and sinks at regional scales are affected by such bias errors, and whether the WMO recommendation should be revised. Propagating this bias uncertainty to flux space showed that a bias consistent with the current recommendation does not deteriorate retrieved fluxes at country scale. However this is not the case if fluxes at a higher spatial resolution (e.g. 0.25°) are of interest. Results showed that observations, although following the WMO recommendations, have a limited ability to detect the interannual flux variability at grid scale. In fact this will be aggravated in future, as atmospheric models are expected to have reduced transport uncertainties leaving space for biases in observations to become dominant.

Zusammenfassung

Langzeitmessungen der Konzentrationen atmosphärischer Treibhausgase an über dem Kontinent verteilten Messstationen verbessern unser Verständnis von Quellen und Senken der Treibhausgase. Diese Messungen können über ein Modell des atmosphärischen Transportes mit den Austauschflüssen zwischen Erdoberfläche und Atmosphäre gekoppelt werden. Innerhalb des EU Projekt ICOS (Integrated Carbon Observation System) wird das Netz der bereits existierenden Messstationen gegenwärtig weiter ausgebaut. Das Design dieses Beobachtungsnetzes erfordert die Maximierung bzw. Optimierung des Informationsgehaltes der atmosphärischen Messungen über die Austauschflüsse betreffende spezifische Zielgrößen. Eine quantitative Netzwerkdesignstudie ist dazu erfordert, mit deren Hilfe das Potential möglicher Konfigurationen von Beobachtungsnetzen abgeschätzt werden kann.

Daher wurde ein regionales Inversionsmodell aufgesetzt, mit dem regionale biogene Austauschflüsse aus CO_2 Messungen an hohen Messtürmen berechnet werden. Das Modellsystem besteht aus den folgenden Komponenten: dem globalen Transportmodell TM3, dem STILT Modell (Stochastic Time-Inverted Lagrangian Transport model), dem diagnostischem VPRM (Vegetation Photosynthesis and Respiration Model) für die a-priori biogenen Austauschflüsse, den Emissionen aus der Verbrennung fossiler Brennstoffe, den Austauschflüssen der Ozeane mit der Atmosphäre, sowie einem Bayes'sches Inversionssystem.

Die Lösung des Inversionsproblems ist unterbestimmt, da nur eine begrenzte Anzahl von Beobachtungen verfügbar sind, das Problem aber deutlich mehr Freiheitsgrade in den zu bestimmenden Unbekannten (zeitlich und räumlich aufgelöste Flüsse) besitzt. Der Bayes'sche Ansatz eröffnet durch die Verwendung von a-priori Flüssen eine effiziente Methode dieses Problem zu lösen. Außerdem ermöglicht er die quantitative Verwendung von Unsicherheiten in sowohl den a-priori Flüssen als auch in den Beobachtungen. Die Bestimmung der optimierten (a posteriori) Flüsse ergibt sich Kombination von atmosphärischen Beobachtungen und a-priori Flüssen, wobei deren Gewichtung den jeweiligen Unsicherheiten entspricht.

Atmosphärische CO_2 Messungen an hohen Türmen sind stark beeinflusst durch die Austauschflüsse an der Erdoberfläche im Umfeld der Station (ungefähr im 100 km Radius). Um die Varianz der Flüsse auf regionaler Skala zu erfassen, die von globalen Modellsystemen nicht mehr erfasst werden können, ist die Verwendung eines hochaufgelösten Systems notwendig. Daher wurde eine hohe räumliche Auflösung (0.25 Grad Auflösung) verwendet, um das atmosphärische Transportmodell mit biogenen Flüssen zu koppeln. Das Ziel der vorliegenden Dissertation ist die quantitative Bestimmung des Einflusses verschiedener Netzwerkkonfigurationen. Dazu wurde eine Analyse der Reduktion der Unsicherheiten vorgenommen. Eine derartige Analyse hängt stark von der a-priori Fehlerstruktur ab. Für eine objektive Netzwerkbewertung sollten daher ausschließlich realistische a-priori Unsicherheiten angenommen werden.

Die quantitative Betrachtung der Unsicherheit ist unerlässlich für die quantitative Bestimmung der biogenen atmosphärischen Austauschflüsse. Um die Unsicherheiten der Flüsse besser in dem Inversionssystem beschreiben zu können, untersucht diese Dissertation zunächst die Fehlerstruktur der Flüsse und erklärt, wie diese Unsicherheiten räumlich und zeitlich korreliert sind. Dabei wurden Flüsse von 3 verschiedenen Biosphären-Modellen mit Messungen an 53

räumlich verteilten Eddy-Kovarianz Messstationen sowie mit Daten aus einer Flugzeugmesskampagne verglichen.

Die räumliche und zeitliche Autokorrelation der täglich aggregierten Residuen der Flüsse (Differenzen der Modellierten und gemessenen Flüsse) kann durch einen exponentiellen Abfall angenähert werden. Die Autokorrelation der täglich aggregierten Modell-Daten Residuen der Flüsse zeigte dabei eine zeitliche Korrelation von 30 Tagen. Für die räumliche Autokorrelation ergaben sich Längenskalen von nur 40 km. Werden jedoch Residuen der Flüsse von jeweils zwei verschiedenen Modellen verwendet, liegen räumliche Autokorrelationslänge im Bereich von einigen 100km. Die Annahme dieser Modell-Daten Fehlerstruktur und eine Hochskalierung auf die europäische Domain führte zu einer jährlich domainweiten Unsicherheit der Kohlenstoffbilanz von 0.15 $GtCy^{-1}$. Diese ist um einen Faktor zwei kleiner ist als die Unsicherheiten, die sich aus einem Ensemble von verschiedenen modellierten Flüssen ergeben. In der Dissertation wurde daher eine Fehlerinflation mit zwei Termen eingeführt. Der erste Term beschreibt die räumlich-zeitliche Struktur der Fehler und der zweite Term repräsentiert einen weitreichenden Bias. Der Bias-Term hat dabei keine zeitliche Variabilität und folgt räumlich einem vordefinierten Flussfeld (z.B. Respirationsfluss).

Diese quantitative datengetriebene Information über die Fehlerstruktur der Flüsse wurde in das Inversionssystem implementiert und intensiv getestet. Ein synthetisches Experiment mit zwei Biosphären-Modellen wurde durchgeführt, wobei ein Biosphären-Modell das a-priori Flussfeld und das andere Biosphären-Modell die als „known-truth“ bezeichneten Flüsse erzeugte. Dieses Experiment ermöglichte die quantitative Beurteilung der Fähigkeit des Inversionssystems, die a-priori biogenen Flüsse auf verschiedenen räumlichen und zeitlichen Skalen zu optimieren. Eine sehr gute Übereinstimmung ergab sich aus den so optimierten und den als „known-truth“ angenommenen Flüssen sowohl bei regionaler als auch nationaler Skala und jährlichen bzw. monatlichen Zeitskalen. Schließlich wurden Inversionen mit realen Messdaten und unterschiedlich angenommener Fehlerstruktur der Flüsse und Untersuchungen zur Sensitivität durchgeführt. Die dabei bestimmten Kohlenstoffsenken liegen im Bereich zwischen 0.23 und 0.38 $GtC\ y^{-1}$. Eine Validierung der monatlich aggregierten Modellflüssen mit Eddy-Kovarianz-Messungen zeigte eine sehr gute Übereinstimmung bei einer Reduzierung des Bias um 52 %, und einer Reduzierung in der Standardabweichung der Residuen um 24 %.

Bei der anschließenden Netzwerkdesignstudie wurden verschiedene Netzwerkkonfigurationen anhand des gegenwärtigen und zukünftigen ICOS Messnetzwerkes mit Hilfe von zwei Inversionssystemen untersucht. Sowohl das Jena Inversionssystem (TM3-STILT) als auch das niederländische Inversionssystem (VUA) ermöglicht eine Reduzierung in den Unsicherheiten der atmosphärischen Landaustauschflüsse. Die räumliche Verteilung der Reduzierung der Unsicherheiten unterscheidet sich zwischen den beiden Inversionssystemen mit einem deutlich stärker lokalisierten Effekt in TM3-STILT. Der Unterschied in der jährlich aggregierten Reduzierung der Unsicherheiten liegt bei 20%, die räumliche Verteilung zwischen verschiedenen Ländern ist jedoch ähnlich. Die erreichte Reduzierung der Unsicherheiten für das bestehende ICOS Netzwerk ist 10.2% bzw. 12.5% für die beiden Inversionssysteme, jeweils bezogen auf das jährliche domainweite Mittel. Für das zukünftige ICOS Netzwerk zeigte sich die Reduzierung der domainweiten Unsicherheiten im Bereich von 11.8% bzw. 17.4 % (für VUA und TM3-STILT).

Atmosphärische Beobachtungen der CO_2 Konzentrationen besitzen eine Messunsicherheit einschließlich eines Bias Fehlers. Für CO_2 empfiehlt die WMO diesen unter 0.1 ppm zu halten. Das Inversionssystem wurde angewendet, um die Auswirkungen eines solchen Bias in den Beobachtungen auf die Bestimmung der Quellen und Senken der Flüsse auf regionaler Skala zu studieren und um einzuordnen, ob die WMO Empfehlung gegeben falls zu revidieren wäre. Dabei zeigte sich, dass der empfohlene Bias-Bereich der Messunsicherheiten für die Bestimmung der aggregierten Flüsse auf nationaler Ebene ausreichend ist. Dies ist jedoch nicht mehr gegeben, wenn Aussagen über Flüsse mit einer hohen räumlichen Auflösung (z.B. 0.25 Grad Auflösung) getroffen werden sollen. Die Resultate zeigen, dass Messungen, die der WMO Empfehlung folgen, dann nur noch eine begrenzte Aussagekraft auf die Flussvariabilität zwischen verschiedenen Jahren auf der regionalen Skala haben. Diese Situation wird sich in der Zukunft verschärfen, wenn durch die Verbesserung der atmosphärischen Modelle eine Reduzierung der Unsicherheit in den atmosphärischen Transportmodellen eintritt und daher der Bias in den Messungen mehr in den Fokus rückt.

Contents

ABSTRACT ..... **I**

ZUSAMMENFASSUNG ..... **III**

CONTENTS ..... **VII**

ACRONYMS ..... **XI**

LIST OF FIGURES ..... **XIII**

LIST OF TABLES ..... **XVII**

CHAPTER 1

1 Introduction 1

1.1 Climate change and the role of carbon dioxide 1

1.2 CO_2 atmospheric inversions: An alternative to bottom-up estimation 3

1.3 European CO_2 inversions and network design 5

1.4 Thesis objectives and outline 6

CHAPTER 2

2 Modeling Framework 9

2.1 Stochastic Time Inverted Lagrangian Transport model (STILT) 9

2.2 Biosphere models 10

2.3 Priors and uncertainties 10

2.4 Jena Inversion System 11

2.5 The experimental framework 12

CHAPTER 3

3 An objective prior error quantification for regional atmospheric inverse applications 15

3.1 Introduction 16

3.2 Data and Methods 18

3.2.1 Observations 19

3.2.2 Biosphere models 23

3.2.3 Analysis of model-observation differences 25
3.2.4 Analysis of model-model differences 27

3.3 Results 28
3.3.1 Model-data comparison for tower and aircraft fluxes 28
3.3.2 Model-model comparison 36

3.4 Discussion and conclusions 38

CHAPTER 4

4 Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. part1: methodology and system evaluation 45

4.1 Introduction 46

4.2 Methods 49
4.2.1 Inversion scheme 49
4.2.2 Characteristics of the inversion set up 51
4.2.2.1 A-priori information and uncertainties 51
4.2.2.2 State space 54
4.2.2.3 Observation vector and uncertainties 54
4.2.2.4 Atmospheric transport 57
4.2.3 Metrics for performance evaluation 57

4.3 Results 58
4.3.1 CO_2 mole fractions 58
4.3.2 Flux estimates and uncertainties 61
4.3.3 Evaluation with synthetic eddy covariance data 67

4.4 Discussion 68
4.4.1 Performance in flux space 68
4.4.2 Performance in observation space 69

4.5 Conclusions 69

CHAPTER 5

5 Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part2: the european terrestrial CO_2 fluxes 73

5.1 Introduction 73

5.2 Methods 74
5.2.1 Two-step inversion 74
5.2.2 Atmospheric network and data 76
5.2.3 A-priori information and uncertainties 79
5.2.4 Diagnostics and aggregation of fluxes 80

5.3 Results 81
5.3.1 Simulated CO_2 and goodness of fit 81
5.3.2 Posterior flux estimates at different scales 86
5.3.3 Validation against eddy covariance measurements 89

5.4 Discussion 91
5.4.1 Goodness of fit 91
5.4.2 Validation against eddy flux measurements 92
5.4.3 Reliability of European flux estimates 93
5.4.3.1 Mismatch in bottom-up and top-down methods 93
5.4.3.2 Sensitivity to anthropogenic emissions 94
5.4.3.3 Sensitivity to site selection 95
5.4.3.4 Retrieved fluxes and comparison to previous inverse estimates 95

5.5 Conclusions 97

CHAPTER 6

6 Assessing the icos current and future atmospheric network using multiple inversion systems 99

6.1 Introduction 99

6.2 Methods 101
6.2.1 Atmospheric inversions and uncertainty reduction as a tool for network design 101
6.2.2 Inversion Systems 103
6.2.2.1 Jena Inversion System TM3-STILT 103
6.2.2.2 Inversion system VUA 103
6.2.3 Joint protocol for regional inversions 104
6.2.3.1 Observational networks 104
6.2.3.2 Prior error setup 108
6.2.3.3 Model-data mismatch error setup 108
6.2.4 Measurement bias experiment 109
6.2.4.1 Setup characteristics 109
6.2.4.2 Targeted uncertainties for flux estimates 110
6.2.4.3 Computation of the flux retrieval bias 110

6.3 Results 111
6.3.1 Assessment of current ICOS network and comparison of the uncer-tainty reduction from different inversion systems 111
6.3.1.1 Uncertainty reduction at grid scale 111
6.3.1.2 Uncertainty reduction at the national scale 114
6.3.1.3 What do we learn from the future ICOS network? 115
6.3.1.4 Impact of a network gap 116
6.3.2 How do biased observations impact flux retrievals? 117

6.4 Discussion 120
6.4.1 Uncertainty reduction metric 120
6.4.2 Measurement bias impact on flux retrievals 121

CHAPTER 7

7 Synthesis of the results – Outlook 125

7.1 On the prior error 126
7.1.1 What is the error structure of the prior error covariance used in atmospheric CO_2 inversions? 126
7.1.2 Can we generalize and use this error structure in every inversion system? 126

7.2 On the regional atmospheric inversions 127
7.2.1 What can we gain from and what are the limitations of a regional high resolution inversion system? 127
7.2.2 What is the biospheric carbon budget in Europe using data driven prior uncertainties and a regional high resolution inversion system? 127

7.3 On the network design 128
7.3.1 How well does the current ICOS network perform, and what will be the gain due to future expansion? 128
7.3.2 Does a bias in atmospheric observations affect flux retrievals, and how accurate should observations be? 128

7.4 Outlook 129

BIBLIOGRAPHY 131

ACKNOWLEDGMENTS 147

SELBSTSTÄNDIGKEITSERKLÄRUNG 149

CURRICULUM VITÆ 151

Acronyms

CCDAS Carbon Cycle Data Assimilation System

CERES Carbo Europe Regional Experiment Strategy

DoI Domain of Interest

ECMWF European Centre for Medium-Range Weather Forecasts

EDGAR Emision Database of Global Atmospheric Research

EVI Enhanced Vegetation Index

GAW Global Atmospheric Watch

GEE Gross Ecosystem Exchange

GHG Green House Gas

HYSPLIT Hybrid Single Particle Lagrangian Integrated Trajectory model

ICOS Integrated Carbon Observing System

LAI Leaf Area Index

LIDAR LIght Detection And Ranging

LPDM Lagrangian Particle Dispersion Model

LSCE Le Laboratoire des Sciences du Climat et de l'Environnement

LSWI Land Surface Water Index

MODIS Moderate Resolution Imaging Spectroradiometer

MPI-BGC Max Planck Institute for BioGeoChemistry

NEE Net Ecosystem Exchange

NOAA National Oceanic and Atmospheric Administration

ORCHIDEE Organizing Carbon and Hydrology In Dynamic Ecosystems

PAR Photosynthetically Available Radiation

PBL Planetary Boundary Layer

PFT Plant Functional Type

RAMS Regional Atmospheric Modelling System

RF Radiative Forcing

STILT Stochastic Time Inverted Langrangian Transport model

SYNMAP Synergetic land cover Product

TM3 Transport Model 3

VPRM Vegetation Photosynthesis and Repsiration Model

VUA Vrije Universiteit Amsterdam

WDCGG World Data Centre for Greenhouse Gases

WMO World Meteorological Organization

5PM 5 Parameter Model

List of Figures

Figure 1.1 Historic and current atmospheric carbon dioxide concentration measured in ice cores and at Mauna Loa station located in Hawaii. CO_2 levels had been stable at about 280 parts per million (ppm) until the beginning of the industrial revolution around 1750. 2

Figure 1.2 Global mean RFs from the agents and mechanisms grouped by agent type. Anthropogenic RFs and the natural direct solar RF are shown. Columns indicate other characteristics of the RF; efficacies are not used to modify the RFs shown. Time scales represent the length of time that a given RF term would persist in the atmosphere after the associated emissions and changes ceased. No CO_2 time scale is given, as its removal from the atmosphere involves a range of processes that can span long time scales, and thus cannot be expressed accurately with a narrow range of lifetime values. Figure is obtained from Forster et al. (2007). 3

Figure 1.3 The GAW (Global Atmosphere Watch) observational network, for CO2 dry mole fractions, provided by the World Meteorological Organization (WMO). Figure available at http://www.wmo.int/pages/prog/arep/gaw/ghg/documents/GHG_Bulletin_No.8_en.pdf. 4

Figure 2.1 Schematic flow chart of the model framework used throughout the thesis. Panel a) describes the real data inversions, using the two-step scheme, and panel b) describes the synthetic inversion. 12

Figure 3.1 Eddy covariance sites used in the study. The dashed line delimits the exact domain used to calculate the aggregated fluxes. 22

Figure 3.2 Box and whisker plot for site-specific correlation coefficients between modeled and observed daily fluxes as a function of the vegetation type. The numbers beneath the x-axis indicate the number of sites involved. The bottom and the top of the box denote the first and the third quartiles. The band inside the box indicates the central 50% and the line within is the median. Upper and lower line edges denote the maximum and the minimum values excluding outliers. Outliers are shown as circles. 29

Figure 3.3 Box and whisker plot for the annual site-specific biases of the models differentiated by vegetation type. Units at y-axis are in μmol m^{-2} s^{-1} (for conversion to gC m^{-2} yr^{-1} reported values in y axis should be multiplied by 378,7694). 30

Figure 3.4 Temporal lagged autocorrelation from model-data daily averaged NEE residuals for all models. Thin red lines correspond to different sites, while the blue thin lines reveal the sites with a bias larger than +/-2.5 μmol m^{-2} s^{-1}. The thick black line shows the all-site autocorrelation, and the thick grey line indicates the all-site autocorrelation but for a sub-set that excludes sites with large model-data bias ("sub-site"). The dark green line is the all-site exponential fit, and the light green line shows the all-site autocorrelation excluding the sites with large bias. The exponential fits use lag times up to 180 days. 31

Figure 3.5 Temporal autocorrelation for VPRM10 – aircraft NEE residuals. Black dots represent individual flux transects pairs sampled at different times as function of time separation. Black circles represent daily scale binned data. 33

Figure 3.6 Distance correlogram for the daily net ecosystem exchange (NEE) residuals using all sites. Black dots represent the different site pairs; the blue line represents the median value of the points per 100-km bin and the green an exponential fit. Results are shown for residuals of VPRM at a resolution of 10 km (top left) and 1 km (top right), ORCHIDEE (bottom left), 5PM (bottom right). 34

Figure 3.7 Annual and seasonal e-folding correlation length of the daily averaged model-data NEE residuals for VPRM at 10 and 1 km resolution, ORCHIDEE and 5PM. "S" refers to the standard case where all pairs were used, "D" refers to the case where only pairs with different vegetation types were used, "I" denotes the case in which only pairs with identical vegetation type were considered, and "*" denotes that in addition 150 days of common non-missing data are required for each pair of sites. The dot represents the best-fit value when fitting the exponential model. The upper and the lower edge of the error bars show the 2.5 and 97.5 percentiles of the length value. Note the scale change in the y-axis at 100 km. 35

Figure 3.8 Distance correlogram between VPRM10 and aircraft NEE measurements. Black dots represent the different aircraft grid points pairs; black circles represent 10 km scale binned data. 36

Figure 3.9 Annual and seasonal e-folding correlation length for an ensemble of daily averaged NEE differences between two models without (filled circle) and with random measurement errors added to the modeled fluxes used as reference (crosses). The symbols represents the best fit value when fitting the exponential model, and the upper and lower edge of the error bars show the 2.5 and 97.5 percentiles of the correlation length. The first acronym at the legend represents the model used as reference and the second the model which was compared with. Note that for the VPRM10/VPRM1 case during spring (with and without random error), the 97.5 percentile of the length value exceeds the y-axis and has a value of 1073, 1626 km respectively. 37

Figure 3.10 Annual e-folding correlation lengths as a function of the factor used for scaling the random measurement error, for all model-model combinations. The black dot-dash lines reveal the range of the spatial correlation lengths generated from the model-data comparisons. 38

Figure 4.1 Domain of the inversions (dashed rectangle). Locations of the atmospheric measurement stations are shown with blue marks. Red stars denote the eddy covariance locations used for flux comparisons at grid scale. 52

Figure 4.2 Monthly data coverage plot for the atmospheric stations used in the regional inversions. Left column shows the code name and the right columns show the station class and the assigned uncertainty in units of ppm. "C" stands for continental sites near the surface, "T" for continental tall towers, "S" for stations near shore, "M" for mountain sites, "MU" for mountain sites with diurnal upslope winds and "UP" for urban pollutant. 56

Figure 4.3 Daily nighttime (23:00-4:00 UTC) averages for prior, true, and posterior CO_2 dry mole fraction time series for the mountain site Schauinsland. Time starts at 1st January 2007. 58

Figure 4.4 Taylor diagram for modeled and measured time-series of CO_2 dry mole fractions. Prior (black), true (green, the perfect match of modeled and true time-series) and the different inversion cases (R0 blue; R1 red) are displayed. Different symbols denote different atmospheric stations. The normalized SD was calculated as the ration of the SD of the modeled time-series to the SD of observations. 60

Figure 4.5 Annual spatial distribution for the prior, true, and posterior biogenic flux estimates for the two synthetic inversions S1 and B1 (top two rows), and flux innovation defined as the difference posterior - prior (bottom row). Fluxes are given in units of $gCy^{-1}m^{-2}$. 61

Figure 4.6 Annual integrated influence for 2007 of the current atmospheric network. Footprint influence is presented in a logarithmic scale and units are in $\log_{10}$[ppm/(μmol/m^2/s)] 62

Figure 4.7 Monthly and annual carbon flux budget, integrated over the European domain. Note that both inversions share the same annual prior uncertainty but monthly uncertainties differ. Blue and red error bars denote the prior uncertainty for the B1 and S1 scenarios respectively. 63

Figure 4.8 Temporal evolution of monthly NEE for selected European countries for the synthetic data inversion. . 65

Figure 4.9 Overview of the model performance summarized in a Taylor diagram. Posterior and prior monthly and country scale aggregated biospheric fluxes are compared against the reference fluxes ("true"). Each line corresponds to a different country. The starting point of each arrow shows prior/reference comparison and the ending point the posterior/reference comparison. Ideally the ending point should coincide with the green point which represents the reference model. 66

Figure 4.10 Mean monthly NEE averaged over the 53 different eddy covariance site locations as reported in Kountouris et al. (2015). A priori (black), true (green), and posterior fluxes for scenarios B1 (blue) and S1 (red) are shown. Units are in $gCm^{-2}day^{-1}$. 67

Figure 4.11 a): Mean spatial correlation of monthly fluxes with true fluxes as function of spatial flux aggregation scale for prior fluxes (grey), and for posterior fluxes from scenarios B1 (blue) and S1 (red). b): Mean spatial correlation of fluxes with true fluxes at 2 deg. spatial resolution as function of temporal flux aggregation scale for prior fluxes (grey), and for posterior fluxes from scenarios B1 (blue) and S1 (red). 68

Figure 5.1 Daily nighttime (23:00-4:00 UTC) averages for prior, true, and posterior CO_2 dry mole fraction time series for the Schauinsland site for the real data inversion. Time starts at 1st January 2007. 81

Figure 5.2 Taylor diagram for modeled and observed time-series of CO_2 dry mole fractions. Prior (black), observed (green, the perfect match of modeled and observed time-series) and the different inversion cases (B1 blue; S1 red) are displayed. Different symbols denote different atmospheric stations. The normalized SD was calculated as the ration of the SD of the modeled time-series to the SD of observations. Gray semi-circles show contours of the standard deviation of the model error. 85

Figure 5.3 Annual biogenic flux spatial distribution (top two rows) and flux innovations (posterior - prior) (bottom two rows) as estimated from the different inversions for the real data case. Units are in $gCy^{-1}m^{-2}$. 86

Figure 5.4 Monthly and annual (panel d) biosphere fluxes integrated over the domain. Panel a) shows B1, B2 and S3 cases, b) S1 and S1a and the c) panel shows S1b and S2 cases. Note that all inversions share the same annual prior uncertainty but monthly prior uncertainties differ. Units are in GtC $month^{-1}$ and GtC y^{-1} for monthly and annual fluxes, respectively..... 87

Figure 5.5 Temporal evolution of prior and posterior monthly NEE for selected European countries. 88

Figure 5.6 Temporal evolution of monthly NEE ($gCm^{-2}day^{-1}$) averaged over all EC sites (top left), excluding crop (top right), and using only crop sites (bottom). Uncertainties (error of the mean monthly NEE) are indicated by the shaded areas. 90

Figure 5.7 Annual European biogenic CO_2 fluxes in $GtCy^{-1}$ for the different inversions and comparison to previous studies. Fluxes are upscaled to the TransCom EU domain. Labels of the references are as follows: Ci : Ciais et al. (2000); Gu : Gurney et al. (2004); Ri : Rivier et al. (2010); Pe : Peylin et al. (2013); Re : Reuter et al. (2014). Periods for the inverted fluxes are given below the flux estimates. 96

Figure 6.1 Site locations and types for the current atmospheric network (ICOS1, blue symbols) and additional stations included in the future network (ICOS2, green symbols). Borders with the red color indicate the stations (within the borders) excluded in the ICOSEG configuration. 105

Figure 6.2 Uncertainty reduction maps for ICOS1 network, for different seasons (MAM: March, April, May; JJA: June, July, August; SON: September, October, November) and for the full year 2007 for TM3-STILT (top panels) and the VUA inversion system (bottom panels). Dashed circles around each station indicate the 100km correlation length assumed in the prior error covariance. 112

Figure 6.3 Same as figure 2 but for the ICOS2 network. 113

Figure 6.4 Dependence of spatial pattern of uncertainty reduction on distance to the closest station for the different inversion systems, seasons, and network configurations (left panel). The same quantity when normalizing to the respective uncertainty reduction at zero distance is presented at the right panel. 113

Figure 6.5 Uncertainty reduction for annually (top) and seasonally aggregated (bottom 3 panels) national biosphere-atmosphere exchange for EU28 countries using the current ICOS (ICOS1) and future (ICOS2) network. Note the different colors for each point denote the number of atmospheric stations within a given country. Countries are ordered according to their area as represented within the model domain. 115

Figure 6.6 Difference in the annual uncertainty reduction between ICOS2 and ICOS2EG; the latter has a gap in the network for Germany. 117

Figure 6.7 Impact on flux bias as calculated from eq.7 (left column) and land areas exceeding the threshold value of 14.2 $gCm^{-2}y^{-1}$ (right column) from 0.1 ppm measurement bias at each station and for different model-data mismatch errors (different rows). Units are given in $gCm^{-2}y^{-1}$. 118

Figure 6.8 Impact of 0.1 ppm CO_2 measurement bias on retrieved fluxes for the domain-wide land area and for each EU28 countries, expressed as a fraction of a priori uncertainty in percent for different assumed model-data mismatch errors. Malta is omitted due to its small areal coverage. Countries are ordered by area. 119

List of Tables

Table 3.1 Eddy covariance sites measuring CO_2 fluxes that were used in the analysis. The land cover classification which is used, is coded as follows; CRO, DCF, EVG, MF, GRA, OSH, SAV for crops, deciduous forest, evergreen forest, mixed forest, grass, shrub and savanna respectively. 19

Table 3.2 Annual temporal autocorrelation times in days, from model-data and model-model residuals. The number within the brackets shows the correlation times when excluding sites with large model-data bias from the analysis.32

Table 4.1 Optimized VPRM parameters SW_0, λ_{SW}, α, β for different vegetation classes[a] 51

Table 4.2 Information on the stations used for the regional inversions. Same network applied for the synthetic, and the real data inversions in Kountouris et al. (2016). In first column the term "type" stands for continuous (C) or flask (F) data. 54

Table 4.3 RMSD (first column in ppm) and correlation coefficients (second column) between known truth and prior/posterior CO_2 dry mole fractions for daily "daytime" or "nighttime" averaged values and for each station. The third column shows χ^2, the normalized dry mole fraction mismatch per degree of freedom for 7-day averaged residuals, as a measure of how well the data were fitted. The format for each station is as follows: RMSD | r^2 | χ^2. . 59

Table 4.4 Performance of the two error structures expressed as the spatial RMSD of the optimized monthly and annual NEE fluxes compared to the truth for the whole domain in 64

Table 5.1 Information on the stations used for the regional inversions. Same network applied for the synthetic, and the real data inversions in Kountouris et al. (2016). In first column the term "type" stands for continuous (C) or flask (F) data. Under "Data origin" WDCGG means "World Data Centre for Greenhouse Gases". 76

Table 5.2 Overview of the inversion scenarios. "Shape" describes the internal structure of the bias component (proportional to respiration R or to Net Ecosystem Exchange NEE), and "Time vary" indicates whether the bias component also has temporal variations or not. The fifth column "Prior" represents the terrestrial model used as prior, and "Correlation shape" describes the functional form used for the spatial prior uncertainty correlation, either exponential (E) or hyperbolic (H). The last column indicates whether the full or the reduced station network was assumed. 78

Table 5.3 RMSD (first column in ppm) and correlation coefficients (second column) between observations and prior/posterior CO_2 dry mole fractions for daily "daytime" or "nighttime" averaged values and for each station. The third column shows χ_c^2, the normalized dry mole fraction mismatch per degree of freedom for 7-day averaged residuals, as a measure of how well the data were fitted. The format for each station is as follows: RMSD | r^2 | χ^2. . 82

Table 5.4 Results from Jackknife delete-1 statistics for VPRM estimated domain-wide NEE for different vegetation classes and for all of the land area. The uncertainty in NEE from all land area was derived assuming independence in the vegetation class specific uncertainties. Note the strong asymmetry between the fraction of land area covered by the different vegetation classes and the number of eddy covariance sites used, indicating over/under representation: for example 8 crop sites represent 51% of the land area, while 15 grassland sites represent 5.6% of the land area of Europe. 93

Table 6.1 Current atmospheric measurement sites (ICOS1) and potential locations for the future expansion of ICOS network (ICOS2). Altitude corresponds to the height above sea level, and M_height is the height above ground. Types of stations are coded as follows: tall tower (T), mountain (M), ocean/coastal (OC), urban polluted (UP) and ground (G). 105

Table 6.2 Model-data uncertainties in ppm for tall towers (T), ground stations (G), ocean/coastal stations (OC), mountain stations (M) and urban polluted sites (UP). Data weight refers to the factor that model-data uncertainties are reduced to mimic future transport model capabilities. 109

Table 6.3 Seasonal and annual domain-wide averaged uncertainty reduction for both inversion systems and networks. 112

Chapter 1

Introduction

1.1 Climate change and the role of carbon dioxide

As climate change becomes one of the biggest challenges facing mankind, this has stopped being just a scientific curiosity but it has instead extended to a global problem, recognized also at a national/political level. With numerous chapters and citations to scientific papers, the new Intergovernmental Panel on Climate Change (IPCC) report (IPCC 2014) illustrates how the climate change will be or has already started impacting our lives. Changing precipitation patterns, weather extremes, global warming, impacts on wild life, reduced food production, and negative economic and health impact are some of the numerous climate change consequences reported in IPCC (2014).

Carbon dioxide (CO_2) is a greenhouse gas which directly affects the Earth's energy budget. The mechanism through which CO_2 warms the atmosphere is the ability to absorb infrared radiation (i.e reemitted radiation from earth's surface), and to radiate this energy isotropically, with half the radiation going back to the surface.

CO_2 has shown to be the most important driver of global warming. Since the industrial revolution the CO_2 emissions released to the atmosphere has risen dramatically (IPCC, 2007) and nowadays the atmospheric burden as increased to about 400 ppm according to the National Oceanic and Atmospheric Administration (NOAA), compared to 280 ppm during preindustrial times. Figure 1.1 shows the CO_2 increase, recorded at Mauna Loa station in Hawaii operated by NOAA. This particular data set is strongly linked with climate change discussion, mainly because of the long data record and the station location where data are being obtained at an altitude of 3400 m in the northern subtropics. Such a remote site as Mauna Loa is considered to be representative for the global CO_2 background composition as it does not suffer from short term variations caused by local vegetation flux or anthropogenic emission events.

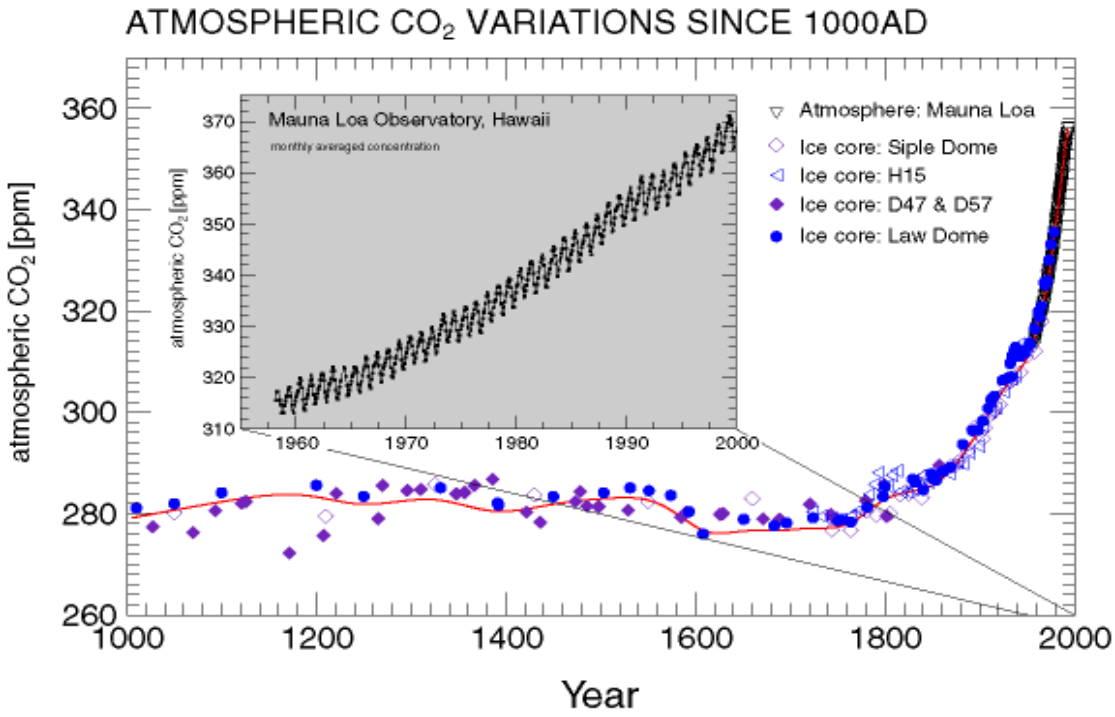

Figure 1.1. Historic and current atmospheric carbon dioxide concentration measured in ice cores and at Mauna Loa station located in Hawaii. CO_2 levels had been stable at about 280 parts per million (ppm) until the beginning of the industrial revolution around 1750.

Human activities also release non-CO_2 greenhouse gases to the atmosphere, causing their abundance to increase significantly. Why does then, CO_2 get most of the attention? IPCC issued a report on global climate change in 2007 that compared the relative influence exerted by key greenhouse gases, aerosols, and land use change of human origin on our climate between 1750 and 2005 (Forster et al., 2007). They use the term "radiative forcing" (RF) which quantifies the disturbances of the earth's radiation balance by a specific atmospheric component (Figure 1.2). As CO_2 has the highest RF (1.66 Wm^{-2}) among all human-influenced climate drivers and considering also the long life time due to its molecular stability (CO_2 does not react with oxygen), it is clear why it gathers the attention of the scientific community.

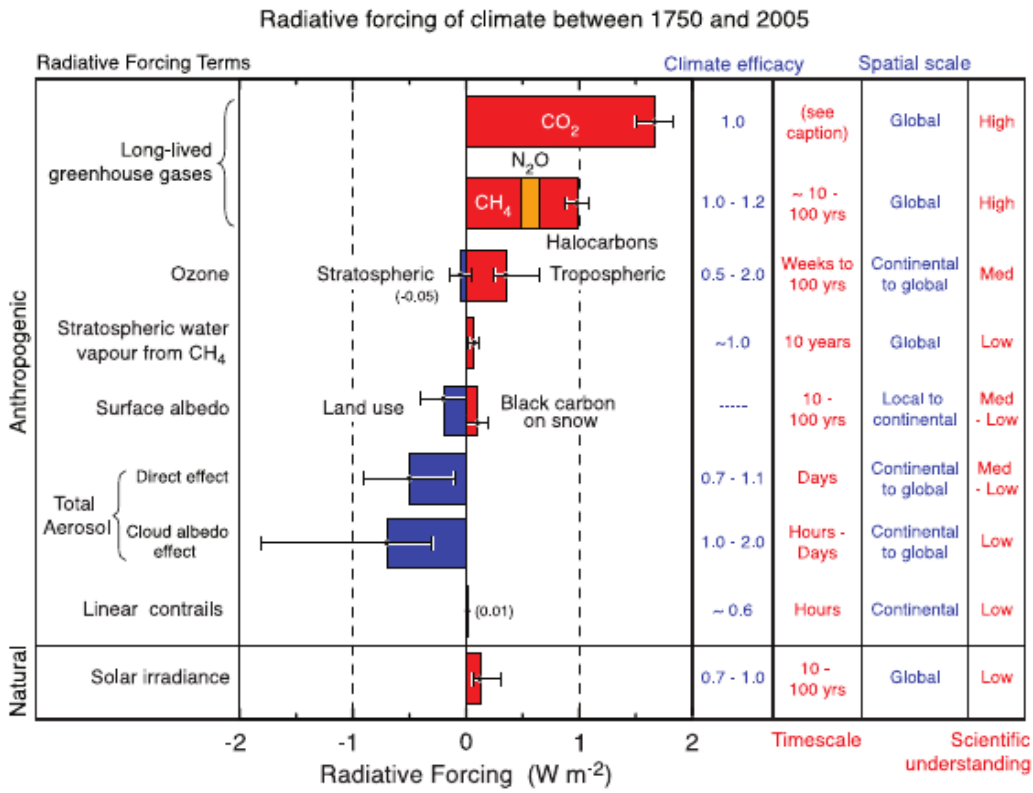

Figure 1.2 Global mean RFs from the agents and mechanisms grouped by agent type. Anthropogenic RFs and the natural direct solar RF are shown. Columns indicate other characteristics of the RF; efficacies are not used to modify the RFs shown. Time scales represent the length of time that a given RF term would persist in the atmosphere after the associated emissions and changes ceased. No CO_2 time scale is given, as its removal from the atmosphere involves a range of processes that can span long time scales, and thus cannot be expressed accurately with a narrow range of lifetime values. Figure is obtained from Forster et al. (2007).

1.2 CO_2 atmospheric inversions: An alternative to bottom-up estimation

Direct measurements of regional carbon fluxes (e.g. eddy covariance technique) are still very sparse to assist at scales that policy making requires. This gap is usually filled by up-scaling flux measurements from local to larger scales or from model intercomparisons but still with large uncertainties. Model intercomparisons show strongly diverse results (Friedlingstein et al., 2006) and further, differences between observations and simulations could reach 10 times the observational uncertainty (Schwalm et al., 2010). Emission inventories are an important source of information regarding anthropogenic CO_2 fluxes. Despite the attempts to accurately record national emissions, it is not only difficult to verify these estimates, but also the differences between such inventories can be as high as 40% (Peylin et al., 2011). An alternative approach to infer CO_2 flux estimates is based on atmospheric dry mole fraction measurements.

Numerous atmospheric stations which measure greenhouse gas dry mole fractions were built over the globe the last decades (see Figure 1.3). As the atmospheric network was continuously expanding, new statistical methods were implemented to make use of the information coming from the observed CO_2 dry mole fractions to infer CO_2 fluxes. In particular, this information can be exploited in an inverse procedure. It makes use of the available observations, an a-priori knowledge of the CO_2 spatiotemporal flux distribution, and an atmospheric transport model, to link carbon sources and sinks to atmospheric observations. Atmospheric inversions are

optimizing the a-priori knowledge or the parameters of an underlying flux model (control or state space) in such a way, that the model – data mismatch of dry mole fractions is minimized. Inversion systems have been extensively used to derive spatiotemporal flux patterns both on global (e.g. Tans et al., 1990; Kaminski et al., 1999a; Gurney et al., 2003; Rödenbeck et al., 2003; Mueller et al., 2008), and on regional scale (Gerbig et al., 2003a; Peylin et al., 2005; Lauvaux et al., 2012; Broquet et al., 2013).

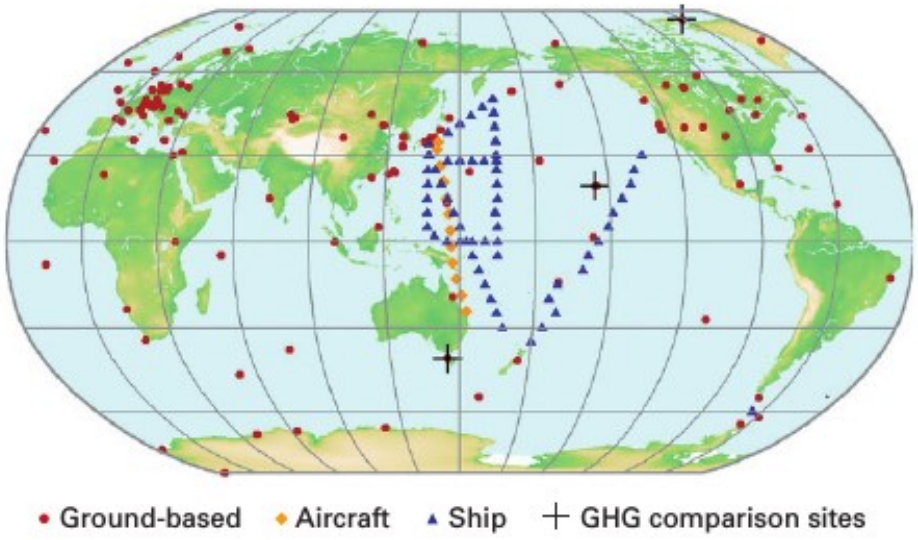

Figure 1.3 The GAW (Global Atmosphere Watch) observational network, for CO2 dry mole fractions, provided by the World Meteorological Organization (WMO). Figure available at http://www.wmo.int/pages/prog/arep/gaw/ghg/documents/GHG_Bulletin_No.8_en.pdf.

One of the most important features in atmospheric inversions is how to deal with the prior uncertainty. Whilst inversion systems are using prior information derived from biosphere models that may be strongly erroneous, an appropriate error structure is needed to optimally weight our trust to the prior information. The prior uncertainty is considered to have a Gaussian distribution, and hence can take the form of a covariance matrix, which includes flux uncertainties at different times and locations. The choice of the error correlations is of critical importance: 1) the higher the correlations, the larger the spatially and temporally aggregated uncertainty will be and 2) they control the spread of the information in space, and in time, from a single observation. Several assumptions have been made in atmospheric inverse studies; they all lack an objective characterization of the error structure. Recently, Chevallier et al. (2006 and 2012) introduced a direct method to characterize the prior error structure from model-data flux mismatches. In those studies, an autocorrelation analysis of the mismatches was used to identify the error correlations. As it was not clear if the error structure they found could be applied to other models apart from ORCHIDEE (Organizing Carbon and Hydrology In Dynamic Ecosystems), we augment their approach in chapter 3, using three biosphere models in comparison to eddy covariance data and further, we study also the model-model autocorrelation structure.

1.3 European CO_2 inversions and network design

Despite the advances in atmospheric inversions, yet global inversions suffer from large uncertainties at regional scale (Baker et al., 2006). A characteristic example are the European fluxes, which are still highly uncertain, because they have been studied mainly by global systems (Ciais et al., 2000; Rödenbeck et al., 2003; Gurney et al., 2004; Peylin et al., 2013). Despite the relatively dense atmospheric observation network that has been developed in Europe, the global systems are not making use of the high temporal resolution (e.g. hourly from continuous stations) of the observations, but rather assimilate daily, weekly or monthly averages (Broquet et al., 2011). Regional scale inversion systems that are using high resolution transport are confirmed to better fit observations, compared to global low resolution transport models (Law et al., 2008; Patra et al., 2008). Until now, studies on biogenic European fluxes at high resolution, and especially for a full year period, are very limited. Peylin et al. (2005) estimated daily CO_2 fluxes over Europe, but only for one month. Broquet et al. (2011) studied also European fluxes at high resolution but only for the summer period. Rivier et al. (2010) used 5 transport models. Among them 3 were regional, but with the highest spatial resolution being at 0.5° x 0.5°, with the rest being 150km x 150km or even coarser. Furthermore, they did not take advantage of the high frequency observations but instead, they used monthly averages. Carouge et al. (2010a and 2010b) focused on the European domain and used a high resolution transport scheme of 40km x 40km. A shortcoming of their studies is that they remained within the frames of a "pseudo-data" experiment, where the observations have been substituted from artificially created ones. This thesis is trying to shed more light on the European biogenic CO_2 fluxes. It uses a regional scale inversion system with a high transport resolution of 0.25° x 0.25°, while it makes use of hourly averaged continuous atmospheric data. All the above is combined with flask measurements from 16 stations across Europe.

In order to better constrain carbon fluxes at high spatial and temporal resolution from atmospheric CO_2 dry mole fractions, a network of monitoring instruments is needed rather than occasional measurements sparse in space and/or in time. Over the last years, a relatively dense network for atmospheric and flux measurements has been established, in the frames of the integrated project CarboEurope-IP. The focus of this 5-year project (started in January 2004) was to quantify and understand the biogenic and oceanic carbon sinks. Currently, through the Integrated Carbon Observing System (ICOS), the atmospheric network is and should be expanded. In detail, the current state of the atmospheric network consists of around 23 stations with the future plan to be around 60 stations (https://www.icos-infrastructure.eu, ICOS Stakeholders Handbook 2013). This endeavor to build such a dense atmospheric network is costly, as new towers are required to be built. To cut down the cost, locations with already existing towers are proposed. To better understand and quantify the additional information from the network expansion, a quantitative network design is needed to assist decision and policy makers.

Network design provides to the decision makers, all the knowledge needed, to determine optimal locations for new stations, as well as assessment of different network configurations. Early studies on network design were carried out by Rayner et al. (1996) and Rayner et al. (2004). In those studies, optimization methods, such as simulated annealing or genetic algorithms were used to construct an optimal (with respect to the location) global network, which minimizes a predefined metric. Lauvaux et al. (2012) assessed two different network configurations, in terms of the number of atmospheric stations used, but only for the North American continent. Lauvaux

et al. (2012) show the importance of network design, since regional carbon fluxes converged to similar annual estimates for both networks, but only the extended network was able to retrieve the spatial flux distribution. Network design studies, focused on the European domain, are very limited. Only recently, a study focusing on network design for the ICOS atmospheric stations was realized by Kadygrov et al. (2015). The shortcoming of this study though is the short time period for which it was applied, considering only a 2-week period in July and December 2007. Further, as transport models differ, it is obvious that network design studies are influenced by the particular transport model that has been used. However, those different results may help quantifying transport uncertainties by using multiple transport models. This thesis performs a network design study on ICOS atmospheric stations, using two different inversion systems (with different transport models and different inversion techniques) for the full year 2007; the results were obtained by deploying the Jena Inversion System, and for the VUA inversion system results were provided by the Vrije University of Amsterdam.

1.4 Thesis objectives and outline

The thesis assesses the ICOS atmospheric observing system, by developing a methodology that is quantitatively tested in every step. Particular emphasis was given to a careful quantification of the prior error structure. The inferred structure was then extensively studied with a state of the art inversion system, focused on inversions at the mesoscale.

This thesis is trying to answer the following scientific questions:

1. What is the error structure of the prior error covariance used in atmospheric CO_2 inversions? Can we generalize and use this structure in every inversion system?

2. What can we gain from, and what are the limitations of a regional high resolution Inversion system? What is the biospheric carbon budget in Europe estimated using data driven prior uncertainties and a regional high resolution inversion system?

3. How well does the current ICOS network perform, and what will be gained from the future expansion? Does a bias in atmospheric observations affect flux retrievals, and how accurate should observations be?

The thesis is organized with the following chapters:

Chapter 2 gives a brief overview of the modeling framework and the key models used in this study.

Chapter 3 presents the method and the analysis performed to describe and quantify the prior error structure. In order to reveal the spatial and temporal autocorrelation structure we studied residuals between three biosphere models and data from eddy covariance flux measurements. This chapter was published as:

Kountouris, P., Gerbig, C., Totsche, K.-U., Dolman, A. J., Meesters, A. G. C. A., Broquet, G., Maignan, F., Gioli, B., Montagnani, L., and Helfter, C.: An objective prior error quantification

for regional atmospheric inverse applications, Biogeosciences, 12, 7403-7421, doi:10.5194/bg-12-7403-2015, 2015.

Chapter 4 introduces the Jena two-step inversion scheme and the set-up of the inversion. In this chapter we evaluate the inversion system in the frame of a synthetic experiment, where the CO_2 fluxes are known and produced by the BIOME-BGCv1 biosphere model. The chapter is based on the manuscript:

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part1: Methodology and system evaluation. Submitted to Atmospheric Chemistry and Physics, acp-2016-577.

Chapter 5 details the results from real data inversions. We used different data-driven prior error structures. Also, sensitivity tests were carried out to infer European carbon fluxes. Flux retrievals were validated with independent data source (eddy covariance flux measurements), and further compared with previous studies. The chapter is based on the manuscript:

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part2: The European terrestrial CO_2 fluxes. Submitted to Atmospheric Chemistry and Physics, acp-2016-578.

Chapter 6 presents a network design study applied to the ICOS atmospheric network, using two inversion systems. The uncertainty reduction analysis is presented here for different temporal and spatial scales. The impact from biased atmospheric measurements on flux retrievals is also analyzed. The chapter is based on the manuscript:

> Kountouris, P., Gerbig, C., Totsche, K.-U., Karstens, U., Koch, F. T., Dolman, A. J., Meesters, A.G.C.A.: Assessing the ICOS current and future atmospheric network using multiple inversion systems. Under preparation.

Finally in **chapter 7**, key findings from this work are summarized, and answer the scientific questions formulated in chapter 1.

Chapter 2

Modeling framework

The modeling framework combines various components. The heart of the system, are the atmospheric transport models TM3 (for global simulations) and STILT (for regional). The transport models are coupled with land and oceanic prior fluxes, as well as with anthropogenic emissions, which are a prerequisite for the Bayesian synthesis inversions. Then, fluxes are optimized through an iterative matrix inversion approach. This chapter briefly introduces the main components used in this work, and outlines the experimental design.

2.1 Stochastic Time Inverted Lagrangian Transport model (STILT)

The Stochastic Time Inverted Lagrangian Transport model (STILT) is a Lagrangian Particle Dispersion Model (LPDM), which is based on the Hybrid Single Particle Lagrangian Integrated Trajectory model (HYSPLIT) (Draxler and Hess, 1998). The STILT model is extensively described in Lin et al. (2003) where it has been used to determine surface fluxes from atmospheric dry mole fractions. It is a receptor oriented model which releases a number of virtual particle ensembles at receptor locations, and calculates their trajectories by driving each particle backward in time. Particles advection is computed using mean wind vectors (ECMWF meteorology) interpolated at particle locations.

The main purpose of using STILT is to derive the sensitivity of the measured CO_2 dry mole fractions to the upstream surface-atmosphere exchange fluxes. This sensitivity, also called footprint *f*, is used as an adjoint of the transport model in the inverse system. The footprint can be calculated as the density ρ of particles at location ($\boldsymbol{x}$,t) that were transported backwards from the receptor location ($\boldsymbol{x}_r$,t_r) normalized by the total number of particles released N_{tot}.

$$f = (x_r, t_r \mid x, t) = \frac{\rho_{part}(x_r, t_r \mid x, t)}{N_{tot}} \qquad 2.1$$

The rate of change in CO_2 dry mole fraction at the receptor point for a given voxel (a value on a regular grid in three-dimensional space, for a time step *m* and a grid-cell i, j of the domain) is the product of the footprint *f* with the actual surface fluxes *F* (Gerbig et al., 2003b; Lin et al., 2003):

$$\Delta C_{m,i,j}(x_r, t_r) = f(x_r, t_r \mid x_i, x_j, t_m) \cdot F(x_i, x_j, t_m) \qquad 2.2$$

STILT has already been used extensively in regional inverse studies and footprint evaluation (Gerbig et al., 2003b; Gerbig et al., 2008; Kort et al., 2008; Göckede et al., 2010; Pillai et al., 2011a).

2.2 Biosphere models

The Vegetation Photosynthesis and Respiration Model (VPRM, Mahadevan et al., 2008) is a diagnostic model, which calculates hourly biosphere-atmosphere CO_2 exchange fluxes; for instance the Net Ecosystem Exchange (NEE). VPRM utilizes meteorological parameters such as the 2-meters temperature, and the short wave radiation. It makes use of satellite observations for the Enhanced Vegetation Index (EVI) and the Land Surface Water Index (LSWI) from MODIS (Moderate Resolution Imaging Spectroradiometer) satellite, which is on board the Terra and Aqua satellites. It uses also the PAR (Photosynthetically Available Radiation), which is the energy flux from the sun in the 400-700 nm range.

VPRM has a simple mathematical structure. Briefly, the model calculates the NEE as the sum of Gross Ecosystem Exchange (GEE) and Respiration (R). The full equation is expressed as

$$NEE = \underbrace{-\gamma \times T_{scale} \times P_{scale} \times W_{scale} (1/[1 + (PAR / PAR_0)]) \times EVI \times PAR}_{GEE} + \underbrace{(\alpha \times T) + \beta}_{R_{eco}} \qquad 2.3$$

where T_{scale}, P_{scale}, and W_{scale} are scalars for temperature, leaf phenology, and canopy water content, respectively. One of the most important input parameters is the vegetation coverage, which is provided by the Synergetic land cover product (SYNMAP) (Jung et al., 2006) at 1 km resolution. The NEE flux calculations are scaled by empirical parameters (i.e. α, β, γ) specific to the vegetation coverage of the model grid cell. These parameters were optimized against eddy covariance flux measurements for the European continent following Mahadevan et al. (2008).

The GBIOME-BGCv1 model (Trusilova and Churkina 2008) estimates daily carbon, nitrogen and water fluxes between the land biosphere and the atmosphere. It is a process-based model and is driven by daily meteorological data, such as maximum and minimum daily temperature, precipitation, vapor pressure deficit, and solar radiation. The land surface is parameterized using a digital elevation map, soil texture data, a land cover classification, including eight plant functional types, atmospheric CO2 concentrations, and the atmospheric deposition of nitrogen. Each plant functional type is described by ecophysiological parameters. More information concerning this model can be found in Trusilova, and Churkina (2008).

Both biosphere models simulate the terrestrial carbon fluxes at a spatial resolution of 0.25º. The VPRM fluxes have hourly temporal resolution, and the GBIOME-BGCv1 fluxes have a daily temporal resolution. However the latter fluxes were disaggregated to hourly resolution; the photosynthetic fluxes were disaggregated by weighting them with the shortwave radiation, whilst hourly respiration fluxes obtained from the daily files simply, by calculating hourly averages (daily value divided by 24).

2.3 Priors and uncertainties

Prior fluxes contain fluxes from the biosphere and they can be derived from a vegetation model. Specifically, in the frames of this thesis the Vegetation Photosynthesis and Respiration Model (VPRM) (Mahadevan et al., 2008) and the BIOME-BGCv1 (Trusilova and Churkina 2008) vegetation models were used. Further, prior fluxes consist also from fossil fuel emissions (taken

from the EDGAR v4.1 inventory[1]), as well as oceanic fluxes (taken from Mikaloff-Fletcher et al. (2007)). EDGAR emissions are scaled at national level for individual years, according to the BP (British Petroleum) statistical review of world energy (BP, 2012), following Steinbach et al. (2011).

Two main types of uncertainty are introduced into the inversion system. The first is the prior uncertainty, expressed in flux space, with units in μmol $m^{-2}s^{-1}$, and represents our trust to the prior fluxes. To quantify the prior error structure, three biosphere models were used (VPRM, 5PM and ORCHIDEE) to simulate daily terrestrial fluxes. Subsequently, observed fluxes from 53 Eddy covariance sites in Europe were compared against the modeled fluxes. Daily flux residuals were investigated for spatial and temporal autocorrelation patterns. Chapter 3 presents in detail the method and the results of this analysis, answering the first research question. The fossil fuel component assumed to be perfectly known thus, no uncertainty assumed.

The second type of uncertainty introduced in the inversion system, is the model-data mismatch, expressed in observation space with units in ppm. This uncertainty type counts for observational errors, uncertainties in transport, and the representation errors.

2.4 Jena Inversion System

Atmospheric inversions are optimizing a-priori flux fields using atmospheric dry mole fractions. Through an iterative procedure fluxes are being adjusted such, that the difference between modeled and observed dry mole fractions is minimized. The modeled dry mole fractions are produced by the inversion system running in forward mode. This thesis uses the Jena Inversion System (Rödenbeck, 2005) and particularly the two-step inversion scheme as described in Trusilova et al. (2010). In the current set-up the two-step scheme is applied to the global transport model TM3 and the regional transport model STILT.

The atmospheric transport models provide the inversion system with the transport operator (also called Jacobian matrix). This matrix links the dry mole fractions at a receptor location to upstream fluxes. TM3 enables a global atmospheric inversion at a coarse horizontal resolution of 5° × 4°, and is being used to create the boundary conditions by performing a forward run, using global optimized fluxes. The boundary conditions represent the influence outside of the domain of interest (DoI), and they are subtracted from the observed dry mole fractions. Afterwards, STILT performs a second, regional inversion for the DoI at high resolution (i.e. 0.5° × 0.5°). More detailed information regarding the Jena Inversion System and the two-step scheme are given in Chapter 4 and in Rödenbeck (2005), Rödenbeck et al. (2009), and Trusilova et al. (2010).

[1]: European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). Emission Database for Global Atmospheric Research (EDGAR), release version 4.1. http://edgar.jrc.ec.europa.eu, 2010.

2.5 The experimental framework

To answer the research questions formulated in chapter 1, we need to establish a model – data framework. A schematic representation of the framework answering the research questions 2 and 3 is given in Figure 2.1.

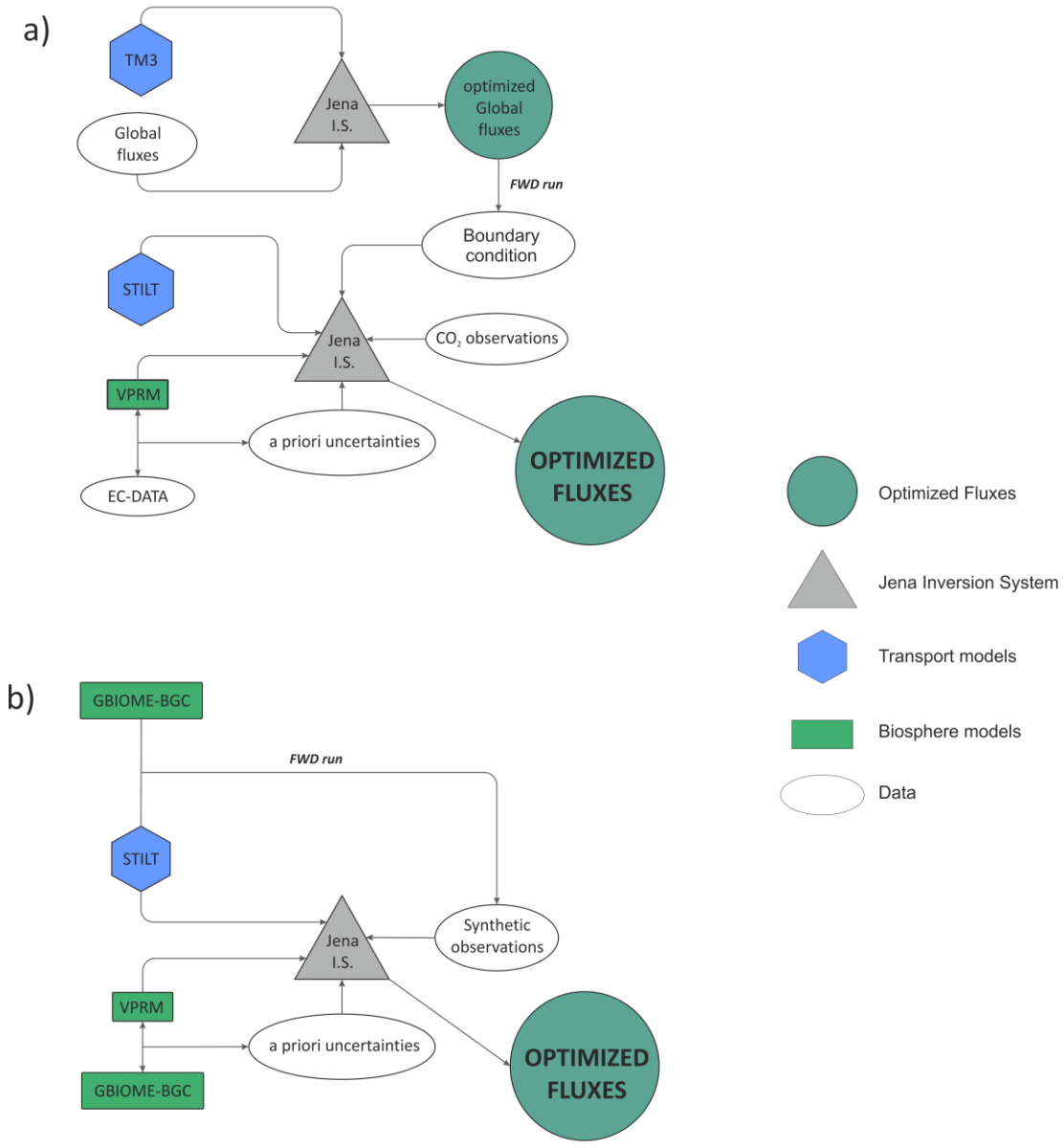

Figure 2.1 Schematic flow chart of the model framework used throughout the thesis. Panel a) describes the real data inversions, using the two-step scheme, and panel b) describes the synthetic inversion.

The two main system components are the transport models TM3 and STILT for simulating the atmospheric transport, while the biosphere models VPRM and GBIOME-BGCv1 are for simulating the biogenic terrestrial fluxes. The transport matrix calculated by the transport models plays a key role in atmospheric inversions. This matrix links the tracer observations (i.e. CO_2) at a receptor point to upstream sources and sinks. Mathematically, this relation between dry mole fractions and upstream fluxes can be expressed as:

$$c = H \cdot f + c_{ini}$$

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} \frac{\partial c_1}{\partial f_1} & \frac{\partial c_1}{\partial f_2} & \cdots & \frac{\partial c_1}{\partial f_m} \\ \frac{\partial c_2}{\partial f_1} & \frac{\partial c_2}{\partial f_2} & \cdots & \frac{\partial c_2}{\partial f_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial c_n}{\partial f_1} & \frac{\partial c_n}{\partial f_2} & \cdots & \frac{\partial c_n}{\partial f_m} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix} + c_{ini} \quad \textbf{2.4}$$

Following this notation, the CO_2 dry mole fractions are represented as a vector c of n observations. This vector contains not only observations at different times, but also at different receptor locations. The transport operator H is a Jacobian matrix, which contains the first order partial derivatives of the flux vector f at a given model grid cell and time. Each row of the H matrix represents the sensitivity of a given dry mole fraction to all m flux elements. In other words, each row represents a footprint calculation from STILT model, at a given location and time of the c_i observation.

A synthetic experiment was designed (lower flow chart in Figure 2.1) in order to evaluate and assess the system performance, after the implementation of the prior error structure. The necessity of using pseudo-data comes from the fact that in the real world we do not have detailed information regarding the spatial and temporal CO_2 flux variability. This experimental framework has two distinct steps. In the first step, a known, true flux field using the biosphere model GBIOME-BGCv1 is created. Assuming a "known and true" flux field, is the best way to evaluate the system, as the correct result is known. Next, a forward run using STILT model is performed to calculate the synthetic observations. In the second step, an atmospheric inversion is realized using a different flux field as prior knowledge, created from VPRM model. Through the inverse procedure VPRM fluxes are optimized to match with those from GBIOME-BGCv1, by minimizing the mismatch in the CO_2 dry mole fractions. The prior error covariance for the synthetic case is calculated by performing an autocorrelation analysis to the residuals between VPRM and GBIOME-BGCv1 models. Using synthetic data, except of the obvious advantage that we create fluxes, which, we perfectly know (and therefore we can directly assess the system performance), provides us with the flexibility to remove uncertainties introduced by the atmospheric transport model. This is achieved because the same transport model is used to generate both the synthetic observations and the inversion itself. Of note is that for the synthetic case boundary conditions are not needed. This happens because the synthetic observations are created only from fluxes within the DoI, and hence they are not influenced from far field fluxes. Therefore, the global model TM3 is not used in this case.

In the real data inversion (upper flow chart in Figure 2.1) observations are affected from far field fluxes and the anthropogenic emissions as well. A global inversion using the TM3 model is realized to optimize the global flux fields. Anthropogenic emissions from EDGAR inventory (version 4.1), were used to represent the atmospheric fossil fuel signal. In the sequel, a forward simulation was performed using the fossil fuel emissions inside the DoI and the global optimized fluxes outside the DoI. The resulted dry mole fractions were subtracted from the observed CO_2 fractions, leaving a pure biospheric only signal. This signal is used to optimize the prior fluxes

inside the DoI at 3-hourly temporal and 0.5° spatial resolution.

In order to answer the third research question, the Jena inversion system was initialized to provide uncertainty reduction estimations at different temporal and spatial scales. The posterior uncertainty estimates are obtained using a Monte Carlo approach. Specifically, 40 error realizations of the prior flux field and the model-data mismatch in the observation space were generated, and an atmospheric inversion was performed for each realization. The posterior flux estimates show a spread from which the posterior flux uncertainties are estimated. The uncertainty reduction is defined as

$$1 - \frac{\sigma_{pr}}{\sigma_{ps}} \qquad \textbf{2.5}$$

where σ are the standard deviations of prior and posterior fluxes across all flux realizations. The uncertainty reduction is calculated at grid, country, and domain-wide scale. As spatial correlations are present in those estimates, aggregating the uncertainty reduction to coarser scale (i.e. country scale) requires that the correlation structure should be retained. Therefore, an aggregation at the desired spatial scale is realized firstly for each of the ensemble members, and secondly then the uncertainty reduction is computed.

Chapter 3

An objective prior error quantification for regional atmospheric inverse applications

Abstract

Assigning proper prior uncertainties for inverse modeling of CO2 is of high importance, both to regularize the otherwise ill-constrained inverse problem, and to quantitatively characterize the magnitude and structure of the error between prior and "true" flux. We use surface fluxes derived from three biosphere models VPRM, ORCHIDEE, and 5PM, and compare them against daily averaged fluxes from 53 Eddy Covariance sites across Europe for the year 2007, and against repeated aircraft flux measurements encompassing spatial transects. In addition we create synthetic observations using modeled fluxes instead of the observed ones, to explore the potential to infer prior uncertainties from model-model residuals. To ensure the realism of the synthetic data analysis, a random measurement noise was added to the modeled tower fluxes which were used as reference. The temporal autocorrelation time for tower model-data residuals was found to be around 30 days for both VPRM and ORCHIDEE, but significantly different for the 5PM model with 70 days. This difference is caused by a few sites with large biases between the data and the 5PM model. The spatial correlation of the model-data residuals for all models was found to be very short, up to few tens of km but with uncertainties up to 100% of this estimation. Propagating this error structure to annual continental – scale yields an uncertainty of 0.06 Gt C and strongly underestimates uncertainties typically used from atmospheric inversion systems, revealing another potential source of errors. Long spatial e-folding correlation lengths up to several hundreds of km were determined when synthetic data were used. Results from repeated aircraft transects in south-western France, are consistent with those obtained from the tower sites in terms of spatial autocorrelation (35 km on average) while temporal autocorrelation is markedly lower (13 days). Our findings suggest that the different prior models have a common temporal error structure. Separating the analysis of the statistics for the model data residuals by seasons did not result in any significant differences of the spatial e-folding correlation lengths.

3.1 Introduction

Atmospheric inversions are widely used to infer surface CO_2 fluxes from observed CO_2 dry mole fractions with a Bayesian approach (Ciais et al., 2000, Gurney et al., 2002, Lauvaux et al., 2008). In this approach a limited number of observations of atmospheric CO_2 mixing ratios are used to solve for generally a much larger number of unknowns, making this an ill-posed problem. By using prior knowledge of the surface-atmosphere exchange fluxes and by using an associated prior uncertainty, the information retrieved in the inversion from the observations is spread out in space and time corresponding to the spatiotemporal structure of the prior uncertainty. In this way, the solution of the otherwise ill-posed problem is regularized in the sense that the optimization problem becomes one with a unique solution. This prior knowledge typically comes from process-oriented or diagnostic biosphere models that simulate the spatiotemporal patterns of terrestrial fluxes, as well as from inventories providing information regarding anthropogenic fluxes such as energy consumption, transportation, industry, and forest fires.

The Bayesian formulation of the inverse problem is a balance between the a priori and the observational constraints. It is crucial to introduce a suitable prior flux field and assign to it proper uncertainties. When prior information is combined with inappropriate prior uncertainties, this can lead to poorly retrieved fluxes (Wu et al., 2011). Here, we are interested in biosphere-atmosphere exchange fluxes and their uncertainties, and make the usual assumption that the uncertainties in anthropogenic emission fluxes are not strongly affecting the atmospheric observations at the rural sites that are used in the regional inversions of biosphere-atmosphere fluxes.

Typically inversions assume that prior uncertainties have a normal and unbiased distribution, and thus can be represented in the form of a covariance matrix. The covariance matrix is a method to weigh our confidence of the prior estimates. The prior error covariance determines to what extent the posterior flux estimates will be constrained by the prior fluxes. Ideally the prior uncertainty should reflect the mismatch between the prior guess and the actual (true) biosphere-atmosphere exchange fluxes. In this sense it needs to also have the corresponding error structure with its spatial and temporal correlations.

A number of different assumptions of the error structure have been considered by atmospheric CO_2 inversion studies. Coarser scale inversions often neglect spatial and temporal correlations as the resolution is low enough for the inverse problem to be regularized (Bousquet et al., 1999, Rödenbeck et al., 2003a) or assume large spatial correlation lengths (several hundreds of km) over land (Houweling et al., 2004, Rödenbeck et al., 2003b). For the former case large correlation scales are implicitly assumed since fluxes within a grid-cell are fully correlated. For regional scale inversions, with higher spatial grid resolutions which are often less than 100 km, the spatial correlations are decreased (Chevallier et al., 2012) and the error structure need to be carefully defined. A variety of different assumptions exist. This is because only recently an objective approach to define prior uncertainties based on mismatch between modeled and observed fluxes has been developed (Chevallier et al., 2006 and 2012). In some regional studies, the same correlations are used as in large scale inversions in order to regularize the problem, although the change of resolution could lead to different correlation scales (Schuh et al., 2010). Alternatively, they are defined with a correlation length representing typical synoptic

meteorological systems (Carouge et al., 2010). In other cases, ad-hoc solutions are adopted, where the correlation lengths are assumed to be smaller than in the case of global inversions (Peylin et al., 2005), or derived from climatological and ecological considerations (Peters et al., 2007) where correlation lengths only within the same ecosystem types have a value of 2000 km. In addition some studies use a number of different correlation structures in order to analyze which seems to be the most appropriate one based on cross-validation of the simulated against observed CO_2 mole fractions. The simulated mole fractions were derived using the influence functions and the inverted fluxes (Lauvaux et al., 2012). Michalak et al., (2004) applied a geostatistical approach based on the Bayesian method, in which the prior probability density function is based on an assumed form of the spatial and temporal correlation and no prior flux estimates are required. It optimizes the prior error covariance parameters, the variance and the spatial correlation length by maximizing the probability density function of the observations with respect to these parameters.

A recent study by Broquet et al. (2013) obtained good agreements between the statistical uncertainties as derived from the inversion system and the actual misfits calculated by comparing the posterior fluxes to local flux measurements at the European and 1-month scale. These good agreements relied in large part on their definition of the prior uncertainties based on the statistics derived in an objective way from model-data mismatch by Chevallier et al., (2006) and Chevallier et al., (2012). In these studies, modeled daily fluxes from a site scale configuration of the ORCHIDEE model are compared with flux observations made within the global FLUXNET site network, based on the eddy covariance method (Baldocchi et al., 2001), and a statistical upscaling technique is used to derive estimates of the uncertainties in ORCHIDEE simulations at lower resolutions. While typical inversion systems have a resolution ranging from tens of kilometers up to several degrees (hundreds of km), with the true resolution of the inverse flux estimates being even coarser, the spatial representativity of the flux observations typically covers an area with a radius of around a kilometer. Considering also the scarcity of the observing sites in the flux network, the spatial information they bring is limited without methods for up-scaling such as the one applied by Chevallier et al., (2012). Typical approaches to up-scale site level fluxes deploy for example model tree algorithms, a machine learning algorithm which is trained to predict carbon flux estimates based on meteorological data, vegetation properties and types (Jung et al., 2009, Xiao et al., 2008), or neural networks (Papale and Valentini 2003). Nevertheless eddy covariance measurements provide a unique opportunity to infer estimates of the prior uncertainties by examining model-data misfits for spatial and temporal autocorrelation structures.

Hilton et al., (2012) studied also the spatial model – data residual error structure using a geostatistical method. Hilton's study is focused on the seasonal scale, i.e. investigated residual errors of seasonally aggregated fluxes. However, the state space (variables to be optimized considering also their temporal resolution) of current inversion systems is often at high temporal resolution (daily or even three-hourly optimizations). Further, the statistical consistency between the error covariance and the state space is crucial. Thus the error structure at the daily time-scale is of interest here, and can be used in atmospheric inversions of the same temporal resolution. Similar to Hilton's study we select an exponentially decaying model to fit the spatial residual autocorrelation.

In this study, we augment the approach of Chevallier et al., (2006 and 2012), to a multi-model - data comparison, investigating among others a potential generalization of the error statistics, suitable to be applied by inversions using different biosphere models as priors. This expectation is derived from the observation that the biosphere models, despite their potential differences typically have much information in common, such as driving meteorological fields, land use maps, or remotely sensed vegetation properties, and sometimes even process descriptions. We evaluate model – model mismatches to (I) investigate intra-model autocorrelation patterns and (II) to explore whether they are consistent with the spatial and temporal e-folding correlation lengths of the model – data mismatch comparisons. Model comparisons have been used in the past to infer the structure of the prior uncertainties. For example, Rödenbeck et al., (2003b) used prior correlation lengths based on statistical analyses of the variations within an ensemble of biospheric models. This approach is to a certain degree questionable, as it is unclear how far the ensemble of models actually can be used as representative of differences between modeled and true fluxes. However, if a relationship between model – data and model – model statistics can be established for a region with dense network of flux observations, it could be used to derive prior error structure also for regions with a less dense observational network.

Moreover, to improve the knowledge of spatial flux error patterns, we make use of a unique set of aircraft fluxes measured on 2-km spatial windows along intensively sampled transects of several tens of km, ideally resolving spatial and temporal variability of ecosystem fluxes across the landscape without the limitation of the flux network with spatial gaps in between measurement locations. Lauvaux et al., (2009) compared results of a regional inversion against measurements of fluxes from aircraft and towers, while this is the first attempt to use aircraft flux measurements to assess spatial and temporal error correlation structures.

This study focuses on the European domain for 2007 (tower data) and 2005 (aircraft data) and uses output from high-resolution biosphere models that have been used for regional inversions. Eddy covariance tower fluxes were derived from the FLUXNET ecosystem network (Baldocchi et al., 2001), while aircraft fluxes were acquired within the CarboEurope Regional Experiment (CERES) in southern France. The methods and basic information regarding the models are summarized in Section 3.2. The results from model-data and model-model comparisons are detailed in Section 3.3. Discussion and conclusions are following in Section 3.4.

3.2 Data and Methods

Appropriate error statistics for the prior error covariance matrix are derived from comparing the output of three biosphere models which are used as priors for regional scale inversions with flux data from the ecosystem network and aircraft. We investigate spatial and temporal autocorrelation structures of the model-data residuals. The temporal autocorrelation is a measure of similarity between residuals at different times but at the same location as a function of the time difference. The spatial autocorrelation refers to the correlation, at a given time, of the model-data residuals at different locations as a function of spatial distance. With this analysis we can formulate and fit an error model such as an exponentially decaying model, which can be directly used in the mesoscale inversion system to describe the prior error covariance.

3.2.1 Observations

A number of tower sites within the European domain, roughly expanding from -12° E to 35° E and 35° N to 61° N (see also Figure 3.1), provide us with direct measurements of CO_2 biospheric fluxes using the eddy covariance technique. This technique computes fluxes from the covariance between vertical wind velocity and CO_2 dry mole fraction (Aubinet et al., 1999). We use Level 3, quality checked, half hourly observations of net ecosystem exchange fluxes (NEE), downloaded from the European Flux Database (www.europe-fluxdata.eu), and listed by site in Table 3.1. Each site is categorized into different vegetation types (Table 3.1).

Table 3.1. Eddy covariance sites measuring CO_2 fluxes that were used in the analysis. The land cover classification which is used, is coded as follows; CRO, DCF, EVG, MF, GRA, OSH, SAV for crops, deciduous forest, evergreen forest, mixed forest, grass, shrub and savanna respectively.

Site code	Site name	Land cover classification	Latitude	Longitude	Citation
BE-Bra	Brasschaat	MF	51.31	4.52	Gielen et al., 2013
BE-Lon	Lonzee	CRO	50.55	4.74	Moureaux et al., 2006
BE-Vie	Vielsalm	MF	50.31	6.00	Aubinet et al., 2001
CH-Cha	Chamau	GRA	47.21	8.41	Zeeman et al., 2010
CH-Dav	Davos	ENF	46.82	9.86	Zweifel et al., 2010
CH-Fru	Frebel	GRA	47.12	8.54	Zeeman et al., 2010
CH-Lae	Laegern	MF	47.48	8.37	Etzold et al., 2010
CH-Oe1	Oensingen grassland	GRA	47.29	7.73	Ammann et al., 2009
CH-Oe2	Oensingen crop	CRO	47.29	7.73	Dietiker et al., 2010
CZ-BK1	Bily Kriz forest	ENF	49.50	18.54	Taufarova et al., 2014
DE-Geb	Gebesee	CRO	51.10	10.91	Kutsch et al., 2010

DE-Gri	Grillenburg	GRA	50.95	13.51	Prescher et al., 2010
DE-Hai	Hainich	DBF	50.79	10.45	Knohl et al., 2003
DE-Kli	Klingenberg	CRO	50.89	13.52	Prescher et al., 2010
DE-Tha	Tharandt	ENF	50.96	13.57	Prescher et al., 2010
DK-Lva	Rimi	GRA	55.68	12.08	Soussana et al., 2007
ES-Agu	Aguamarga	OSH	36.94	-2.03	Rey et al., 2012
ES-ES2	El Saler-Sueca (Valencia)	CRO	39.28	-0.32	-
ES-LMa	Las Majadas del Tietar (Caceres)	SAV	39.94	-5.77	Casals et al., 2011
FI-Hyy	Hyytiälä	ENF	61.85	24.30	Suni et al., 2003
FR-Aur	Aurade□	CRO	43.55	1.11	Tallec et al., 2013
FR-Avi	Avignon	CRO	43.92	4.88	Garrigues et al., 2014
FR-Fon	Fontainebleau	DBF	48.48	2.78	Delpierre et al., 2009
FR-Hes	Hesse	DBF	48.67	7.07	Longdoz et al., 2008
FR-LBr	Le Bray	ENF	44.72	-0.77	Jarosz el al., 2008
FR-Lq1	Laqueuille intensive	GRA	45.64	2.74	Klumpp et al., 2011
FR-Lq2	Laqueuille extensive	GRA	45.64	2.74	Klumpp et al., 2011
FR-Mau	Mauzac	GRA	43.39	1.29	Albergel et al., 2010
FR-Pue	Puechabon	EBF	43.74	3.60	Allard et al., 2008

HU-Mat	Matra	CRO	47.85	19.73	Nagy et al., 2007
IT-Amp	Amplero	GRA	41.90	13.61	Barcza et al., 2007
IT-BCi	Borgo Cioffi	CRO	40.52	14.96	Kutsch et al., 2010
IT-Cas	Castellaro	CRO	45.07	8.72	Meijide et al., 2011
IT-Col	Collelongo	DBF	41.85	13.59	Guidolotti et al., 2013
IT-Cpz	Castelporziano	EBF	41.71	12.38	Garbulsky et al., 2008
IT-Lav	Lavarone	ENF	45.96	11.28	Marcolla et al., 2003
IT-Lec	Lecceto	EBF	43.30	11.27	Chiesi et al., 2011
IT-LMa	Malga Arpaco	GRA	46.11	11.70	Soussana et al., 2007
IT-MBo	Monte Bondone	GRA	46.01	11.05	Marcolla et al., 2011
IT-Ren	Renon	ENF	46.59	11.43	Marcolla et al., 2005
IT-Ro2	Roccarespampani 2	DBF	42.39	11.92	Wei et al., 2014
IT-SRo	San Rossore	ENF	43.73	10.28	Matteucci et al., 2014
NL-Dij	Dijkgraaf	CRO	51.99	5.65	Jans et al., 2010
NL-Loo	Loobos	ENF	52.17	5.74	Elbers et al., 2011
NL-Lut	Lutjewad	CRO	53.40	6.36	Moors et al., 2010
PT-Esp	Espirra	EBF	38.64	-8.60	Gabriel et al., 2013
PT-Mi2	Mitra IV (Tojal)	GRA	38.48	-8.02	Jongen et al., 2011
SE-Kno	KnottŒsen	ENF	61.00	16.22	-

SE-Nor	Norunda	ENF	60.09	17.48	-
SE-Sk1	Skyttorp 1	ENF	60.13	17.92	-
SK-Tat	Tatra	ENF	49.12	20.16	-
	Auchencorth				Helfter et al., 2015
UK-AMo	Moss	GRA	55.79	-3.24	
UK-EBu	Easter Bush	GRA	55.87	-3.21	Skiba et al., 2013

A land cover classification is used to label the sites as crop (17 sites), deciduous forest (6), evergreen forest (17), grassland (8), mixed forest (3), savannah (1 site), and shrub land (1). For the current study we focus on observations from these 53 European sites during the year 2007 (Figure 3.1).

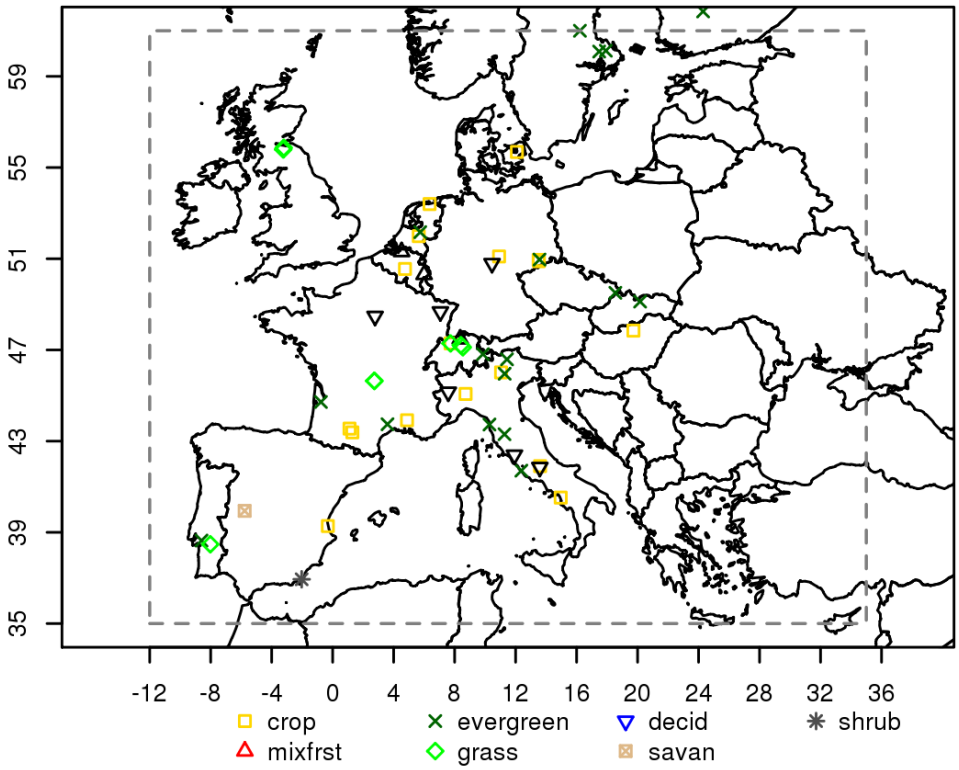

Figure 3.1 Eddy covariance sites used in the study. The dashed line delimits the exact domain used to calculate the aggregated fluxes.

Additionally, aircraft fluxes are used, obtained with an eddy covariance system installed onboard a SkyArrow ERA aircraft (Gioli et al., 2006). Flights were made in southern France during CERES (CarboEurope Regional Experiment) from May 17 to June 22, 2005. Eddy covariance

fluxes were computed on 2-km length spatial windows along transects of 69-km above forest and 78-km above agricultural land, flown 52 and 54 times respectively, covering the daily course. Exact routes are reported in Dolman et al., 2006.

3.2.2 Biosphere models

We simulate CO_2 terrestrial fluxes for 2007 with three different biosphere models described in the following. The "Vegetation Photosynthesis and Respiration Model" (VPRM) (Mahadevan et al., 2008), used to produce prior flux fields for inverse studies (Pillai et al., 2012), is a diagnostic model that uses EVI - enhanced vegetation index and LSWI – land surface water index from MODIS, a vegetation map (Synmap, Jung et al., 2006) and meteorological data (temperature at 2m and downward shortwave radiative flux extracted from ECMWF short term forecast fields at 0.25 degrees resolution) to derive gross biogenic fluxes. VPRM parameters controlling respiration and photosynthesis for different vegetation types (a total of four parameters per vegetation type) were optimized using eddy covariance data for the year 2005 collected during the CarboEuropeIP project (Pillai et al., 2012). For this study, VPRM fluxes are provided at hourly temporal resolution and at three spatial resolutions of 1, 10 and 50 km (referred to as VPRM1,VPRM10 and VPRM50). The difference between the 1,10 and 50 km resolution version is the aggregation of MODIS indices to either 1, 10 or 50 km, otherwise the same meteorology and VPRM parameters are used. At 10 km resolution VPRM uses a tiled approach, with fractional coverage for the different vegetation types, and vegetation type specific values for MODIS indices. For the comparison with the aircraft data VPRM produced fluxes for 2005 at 10 km spatial resolution.

The "Organizing Carbon and Hydrology In Dynamic Ecosystems", ORCHIDEE, model (Krinner et al., 2005) is a process based site scale to global land surface model that simulates the water and carbon cycle using meteorological forcing (temperature, precipitation, humidity, wind, radiation, pressure). The water balance is solved at a half-hourly time step while the main carbon processes (computation of a prognostic LAI, allocation, respiration, turnover) are called on a daily basis. It uses a tiled approach, with fractional coverage for 13 Plant Functional Types (PFT). It has been extensively used as prior information in regional and global scale inversions (Piao et al., 2009, Broquet et al., 2013). For the present simulation, we use a global configuration of the version 1.9.6 of ORCHIDEE, where no parameter has been optimized against eddy covariance data. The model is forced with 0.5° WFDEI meteorological fields (Weedon et al., 2014). The PFT map is derived from an Olson land cover map (Olson 1994) based on AVHRR remote sensing data (Eidenshink and Faundeen 1994). The fluxes are diagnosed at 3-hourly temporal resolution and at 0.5 degree horizontal resolution.

The "5 parameter model" (5PM) (Groenendijk et al., 2011), also used in atmospheric inversions (Tolk et al., 2011, Meesters et al., 2012), is a physiological model describing transpiration, photosynthesis, and respiration. It uses MODIS LAI (leaf area index) at 10km resolution, meteorological data (temperature, moisture, and downward shortwave radiative flux, presently from ECMWF at 0.25 degrees resolution), and differentiates PFTs for different vegetation types and climate regions. 5PM fluxes are at hourly temporal resolution. The optimization has been done with EC-data from Fluxnet as described (except for heterotrophic respiration) in

Groenendijk et al., 2011. Regarding the heterotrophic respiration, an ad hoc optimization using Fluxnet EC-data from 2007 was performed since no previous optimization was available.

Modeled fluxes for all above mentioned sites have been provided by the different models by extracting the fluxes from the grid cells which encompass the EC station location using vegetation type specific simulated fluxes, i.e. using the vegetation type within the respective grid cell for which the eddy covariance site is assumed representative. For most of the sites the same vegetation type was used for model extraction as long as this vegetation type is represented within the grid-cell. As VPRM uses a tile approach, for two cases ("IT-Amp", "IT-MBo") the represented vegetation type (crop) differ from the actual one (grass). For these cases, the fluxes corresponding to crop were extracted. Fluxes were aggregated to daily fluxes in the following way: first, fluxes from VPRM and 5PM as well as the observed fluxes were temporally aggregated to match with the ORCHIDEE 3-hourly resolution; in a second step we created gaps in the modeled fluxes where no observations were available; the last step aggregated to daily resolution on the premise that a) the gaps covered less than 50% of the day, and b) the number of gaps (number of individual 3-hourly missing values) during day and during night were similar (not different by more than a factor two) to avoid biasing.

Spatial and temporal correlation structures and the standard deviation of flux residuals (model-observations) were examined for daily fluxes over the year 2007. Simulated fluxes from the different models are at different spatial resolution, which makes comparisons difficult to interpret. For the model-data residual analysis, the models VPRM1, VPRM10, ORCHIDEE and 5PM were used. We note that VPRM1 with 1 km resolution is considered compatible when comparing with local measurements. For the model-model analysis we use VPRM50 at 50km resolution when comparing with ORCHIDEE fluxes as both models share the same resolution. VPRM10 is considered also appropriate for comparisons with 5PM model as they both share same resolution (MODIS radiation resolution of 1 km aggregated to 10 km and meteorological resolution at 0.25 degrees). Following we compare VPRM50 with 5PM to investigate if the different spatial resolution influences the correlation scale as a measure of how trustful might be the derived scales from ORCHIDEE – 5PM comparisons.

For the aircraft analysis, only the VPRM was used since it is the only model with spatial resolution (10 km) comparable with aircraft flux footprint and capable of resolving spatial variability in relatively short flight distances. Aircraft NEE data, natively at 2 km resolution along the track, have been aggregated into 10 km segments, to maximize the overlap with the VPRM grid, obtaining 6 grid points in forest transects and 8 in agricultural land transects. Footprint areas of aircraft fluxes were computed with the analytical model of Hsieh et al. (2000), yielding an average footprint width containing 90% of the flux of 3.9 km. Averaging also over the different wind directions (perpendicular or parallel to the flight direction), and taking into account the 10 km length of the segments, the area that the aircraft flux data corresponds to, is arround 23.5 km ± 12 km^2. VPRM fluxes at each aircraft grid cell were extracted, and then linearly interpolated to the time of each flux observation.

3.2.3 Analysis of model-observation differences

Observed and modeled fluxes are represented as the sum of the measured or simulated values and an error term, respectively. When we compare modeled to observed data this error term is a combination of model (the prior uncertainty we are interested in) and observation error. Separating the observation error from the model error in the statistical analysis of the model-observation mismatch is not possible; therefore e-folding correlation length estimations do include the observation error term. Nevertheless later in the analysis of model-model differences we assess the impact of the observation error on estimated e-folding correlation lengths.

The tower temporal autocorrelation is computed between the time series of model-observations differences $x_{l,i}$ at site l and the same series lagged by a time unit k (Eq. 3.1), where $\overline{x}$ is the overall mean and N the number of observations:

$$r_l(k) = \frac{\sum_{i=1}^{N-k} (x_{l,i} - \overline{x_l}) \cdot (x_{l,i+k} - \overline{x_l})}{\sum_{i=1}^{N} (x_{l,i} - \overline{x_l})^2}$$

3.1

In order to reduce boundary effects in the computation of the autocorrelation at lag times around one year, the one-year flux time series data (model and observations) for each site was replicated four times. This follows the approach of Chevallier et al., (2012), where sites with at least three consecutive years of measurements have been used.

In the current analysis we introduce the all-site temporal autocorrelation by simultaneously computing the autocorrelation for all the observation sites, with M the number of the sites:

$$r(k) = \frac{\sum_{l=1}^{M} \sum_{i=1}^{N-k} (x_{l,i} - \overline{x_l}) \cdot (x_{l,i+k} - \overline{x_l})}{\sum_{l=1}^{M} \sum_{i=1}^{N} (x_{l,i} - \overline{x_l})^2}$$

3.2

Temporal correlation scales τ were derived by fitting an exponentially decaying model:

$$r = (1 - \alpha) \cdot e^{-\frac{t}{\tau}}$$

3.3

Here t is the time lag. For the exponential fit, lags up to 180 days were used (thus the increase in correlations for lag times larger than 10 months is excluded). At zero lag time the correlogram has a value of one (fully correlated), however for even small lag times this drops to values

smaller than one, also known as the nugget effect. The nugget effect is driven by measurement errors and variations at distances (spatial or temporal) smaller than the sampling interval. For this we include the nugget effect variable α.

The aircraft temporal autocorrelation was similarly computed according to Eq. 3.1 using VPRM, and the same exponentially decaying model (Eq. 3.3) was used to fit the individual flight flux data. The temporal interval was limited at 36 days by the experiment duration.

For the spatial analysis the correlation between model-observation residuals at two different locations (i.e sites or aircraft grid points) separated by a specific distance was computed in a way similar to the temporal correlation, and involved all possible pairs of sites and aircraft grid points. Additional data treatment for the spatial analysis was applied to reduce the impact of tower data gaps, as it is possible that the time series for two sites might have missing data at different times. Thus in order to have more robust results, we also examined spatial structures by setting a minimum threshold of 150 days of overlapping observations within each site pair. Furthermore spatial correlation was investigated for seasonal dependence, where seasons are defined as summer (JJA), fall (SON), winter (DJF for the same year), and spring (MAM). In those cases a different threshold of 20 days of overlapping observations was applied. We note that we do not intend to investigate the errors at the seasonal scale but rather to study if different seasons trigger different error correlation structures.

To estimate the spatial correlation scales, the pairwise correlations were grouped into bins of 100 km distance for towers and 10 km for aircraft data, respectively (*dist*). Following the median for each bin was calculated, and a model similar to Eq. 3.3 was fitted, but omitting the nugget effect variable:

$$r = e^{-\frac{dist}{d}} \qquad \textbf{3.4}$$

The nugget effect could not be constrained simultaneously with the spatial correlation scale d, given the relatively coarse distance groups, the fast drop in the median correlation from one at zero distance to small values for the first distance bin combined with somewhat variations at larger distances. Note that this difference between the spatial and the temporal correlation becomes obvious in the results section 3.

Confidence intervals for the estimated model parameters were computed based on the profile likelihood (Venzon and Moolgavkar, 1987) as implemented within the “confint” function from MASS package inside the R statistical language.

As aircraft fluxes cannot obviously be measured at the same time at different locations, given the relatively short flight duration (about one hour) we treated aircraft flux transect as instantaneous ‘snapshots’ of the flux spatial pattern across a landscape, neglecting temporal variability that may have occurred during flight.

3.2.4 Analysis of model-model differences

We evaluate both model-data flux residuals and model-model differences in a sense of pairwise model comparisons, in order to assess if model-model differences can be used as proxy for the prior uncertainty, assuming that models have independent prior errors. In order to minimize potential influence of the different spatial resolution between the models on the estimated correlation lengths, we compare pairs that have comparable spatial resolution. Such cases are VPRM50-ORCHIDEE and VPRM10-5PM. We choose VPRM10 as more representative to compare against 5PM as 5PM fluxes have also a resolution of 10 km (the main driver, MODIS radiation has the same resolution of 10 km for both models). Similar to the model-observation analysis, the statistical analysis gives a combined effect of both model errors. We assess the impact in the error structure between model-observation and model-model comparisons caused by the observation error by adding a random measurement error to each model-model comparison. This error has the same characteristics as the observation error which is typically associated with eddy covariance observations; the error characteristics were derived from the paired observation approach (Richardson et al., 2008). Specifically, we implement the flux observation error as a random process (white noise) with a double-exponential probability density function. This can be achieved by selecting a random variable u drawn from the uniform distribution in the interval (-1/2, 1/2), and then applying Eq. 3.5 to get a Laplace distribution (also referred to as the double-exponential)

$$x = \mu - \frac{\sigma}{\sqrt{2}} \cdot \mathrm{sgn}(u) \cdot \ln\left(1 - 2 \cdot |u|\right) \tag{3.5}$$

Here μ=0 and σ is the standard deviation of the double-exponential. We compute the σ according to Richardson et al., (2006) as

$$\sigma = \alpha_1 + \alpha_2 \cdot |F| \tag{3.6}$$

where F is the flux and α_1, α_2 are scalars specific to the different vegetation classes. Lasslop et al., (2008) found that the autocorrelation of the half hourly random errors is below 0.7 for a lag of 30 min, and falls off rapidly for longer lag times. Thus we assume the standard deviation for hourly random errors to be comparable with the half hourly errors. Hourly random errors specific for each reference model are generated for each site individually. With ORCHIDEE as reference with fluxes at 3-hourly resolution, a new ensemble of 3-hourly random noise was generated with σ for the 3-hourly errors modified (divided by the square root of three to be coherent with the hourly σ). As both modeled and observed fluxes share the same gaps, the random errors were aggregated to daily resolution, with gaps such to match those of observed fluxes. Finally the daily random errors were added to the modeled fluxes.

3.3 Results

3.3.1 Model-data comparison for tower and aircraft fluxes

Observed daily averaged NEE fluxes, for all ground sites and the full time-series, yield a standard deviation of 3.01 μmol m^{-2} s^{-1}, while the modeled fluxes were found to be less spatially varying and with a standard deviation of 2.84, 2.80, 2.53, 2.64 μmol m^{-2} s^{-1} for VPRM10, VPRM1, ORCHIDEE and 5PM respectively.

The residual distribution of the models defined as the difference between simulated and observed daily flux averages for the full year 2007 was found to have a standard deviation of 2.47, 2.49, 2.7 and 2.25 μmol m^{-2} s^{-1} for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. Those values are only slightly smaller than the standard deviations of the observed or modeled fluxes themselves. This fact is in line with the generally low fraction of explained variance with r-square values of 0.31, 0.27, 0.12, and 0.25 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. When using site-specific correlations (correlations computed for each site, then averaged over all sites), the average fraction of explained variance increases to 0.38, 0.36, 0.35, and 0.42, for VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. Note that for deseasonalized time-series (using a 2^{nd} order harmonic, not shown) the same picture emerges with increased averaged site specific correlation compared to correlations using all sites. This indicates better performance for the models to simulate temporal changes (not only seasonal, but also synoptic) at the site level. Further, the differences between site-specific to the overall r-square values indicate limitation of the models to reproduce observed spatial (site to site) differences. Figure 3.2 shows the correlation between modeled and observed daily fluxes as a function of the vegetation type characterizing each site. All models exhibit a significant scatter of the correlation ranging from 0.9 for some sites to 0 or even negative correlation for some crop sites, with the highest correlation coefficients for deciduous and mixed forest.

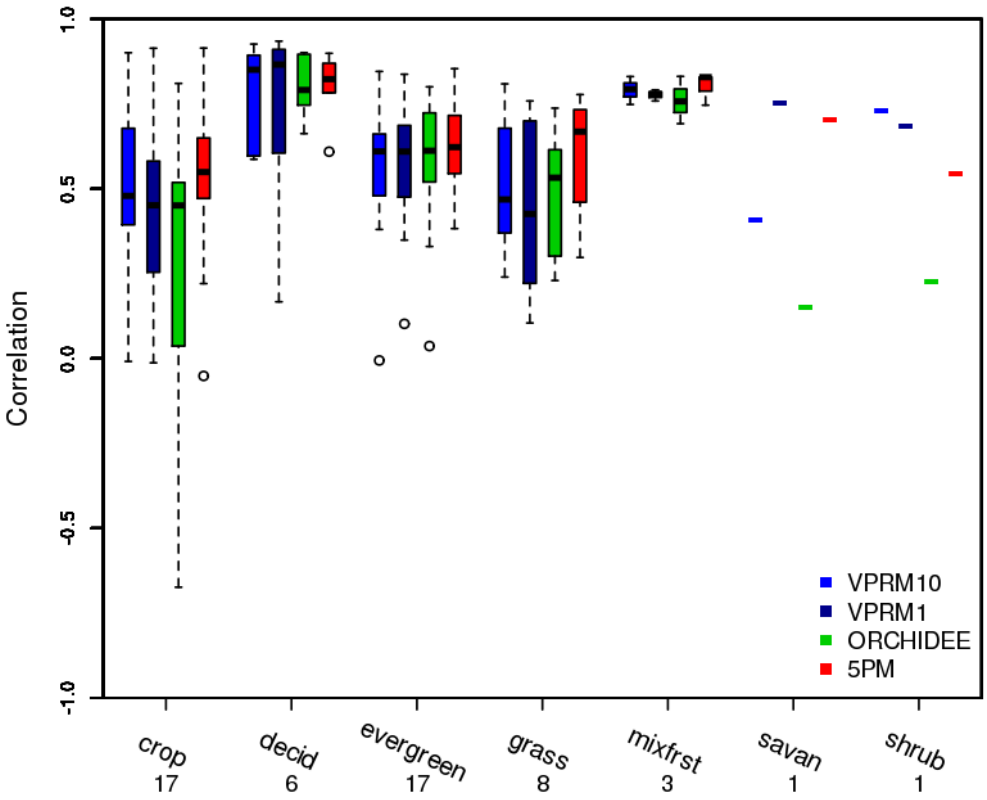

Figure 3.2 Box and whisker plot for site-specific correlation coefficients between modeled and observed daily fluxes as a function of the vegetation type. The numbers beneath the x-axis indicate the number of sites involved. The bottom and the top of the box denote the first and the third quartiles. The band inside the box indicates the central 50% and the line within is the median. Upper and lower line edges denote the maximum and the minimum values excluding outliers. Outliers are shown as circles.

The distribution is biased by -0.07, 0.26, 0.92 and 0.25 μmol m^{-2} s^{-1} for VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. Figure 3.3 shows the distribution of bias (defined as modeled – observed fluxes) for different vegetation types. Bias and standard deviation seem to depend on the vegetation type for all models, without a clear general pattern.

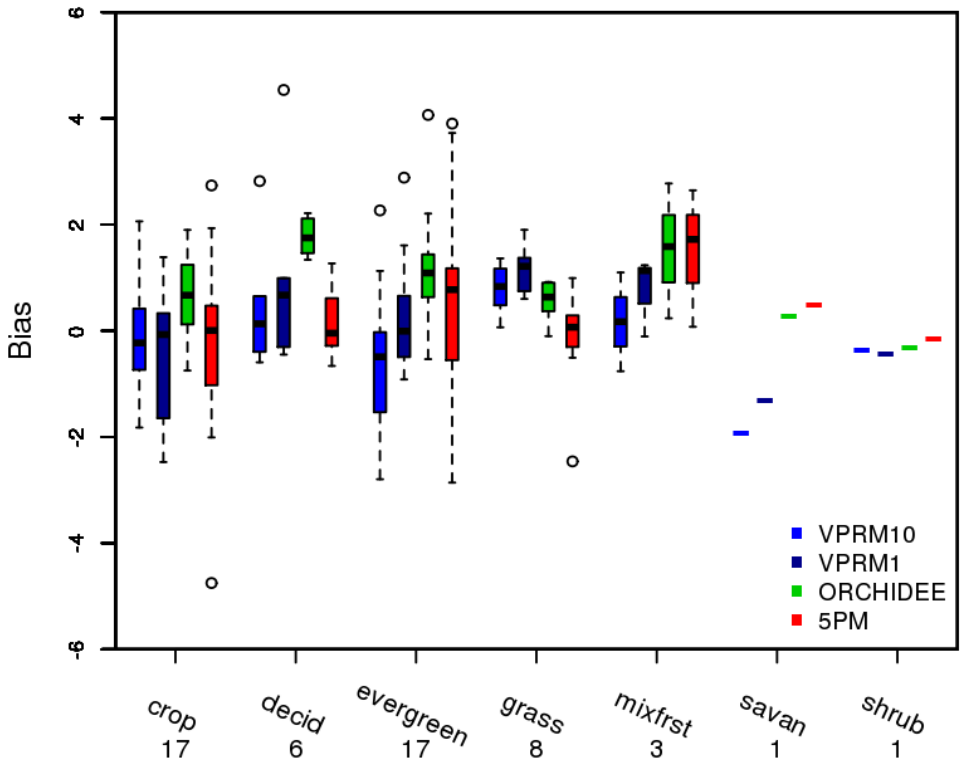

Figure 3.3 Box and whisker plot for the annual site-specific biases of the models differentiated by vegetation type. Units at y-axis are in μmol m^{-2} s^{-1} (for conversion to gC m^{-2} yr^{-1} reported values in y axis should be multiplied by 378,7694).

The temporal autocorrelation was calculated for model-data residuals for each of the flux sites ("site data" in Figure 3.4), but also for the full dataset ("all-site" in Figure 3.4). The "all site" temporal autocorrelation structure of the residuals appears to have the same pattern for all models. It decays smoothly for time lags up to 3 months and then remains constant near to 0 or to some small negative values. The temporal autocorrelation increases again for time lags > 10 months, which is caused by the seasonal cycle. These temporal autocorrelation results agree with the findings of Chevallier et al., (2012).

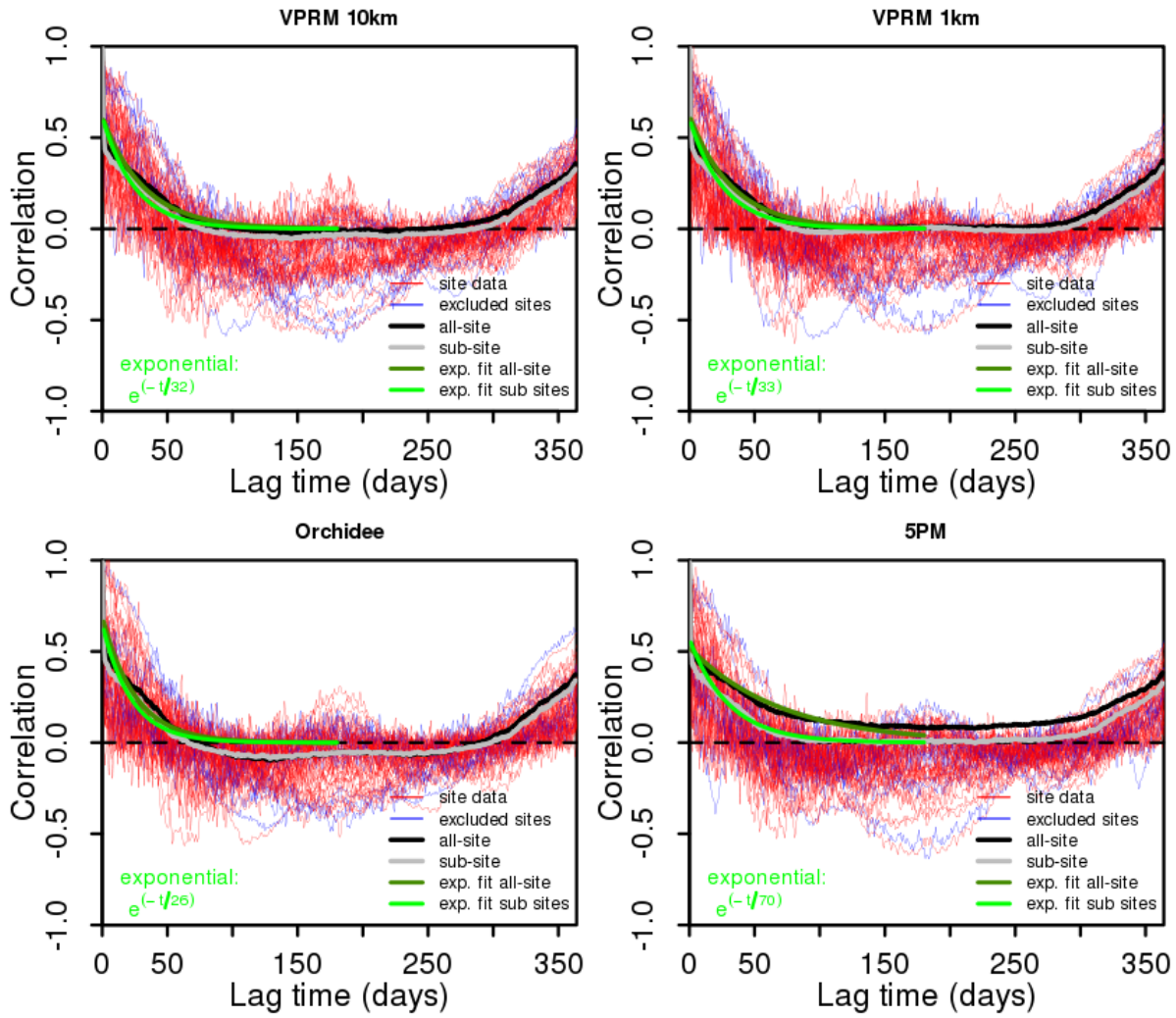

Figure 3.4 Temporal lagged autocorrelation from model-data daily averaged NEE residuals for all models. Thin red lines correspond to different sites, while the blue thin lines reveal the sites with a bias larger than +/-2.5 µmol m^{-2} s^{-1}. The thick black line shows the all-site autocorrelation, and the thick grey line indicates the all-site autocorrelation but for a sub-set that excludes sites with large model-data bias ("sub-site"). The dark green line is the all-site exponential fit, and the light green line shows the all-site autocorrelation excluding the sites with large bias. The exponential fits use lag times up to 180 days.

The exponentially decaying model in Eq. 3.5 was used to fit the data. At zero separation time (t=0) the correlogram value is 1. However the correlogram exhibits a nugget effect (values ranging from 0.31 to 0.48 for the different models) as a consequence of an uncorrelated part of the error. For the current analyses we fit the exponential model with an initial correlation different from 1. The fit has a root mean square error ranging from 0.036 to 0.059 for the different biosphere models. The normalized RMSE (i.e. RMSE divided by the range of the autocorrelation) results in values ranging from 0.061 to 0.092 indicating relative errors in the fit of less than 10%. The e-folding time (defined as the lag required for the correlation to decrease by a factor of e (63% of its initial value) ranged between 26-70 days for the different models (see Table 3.2).

Table 3.2 Annual temporal autocorrelation times in days, from model-data and model-model residuals. The number within the brackets shows the correlation times when excluding sites with large model-data bias from the analysis.

Reference	VPRM10 [days]	VPRM1 [days]	ORCHIDEE [days]	5PM [days]
OBSERVATION	32 (27)	33 (29)	26 (24)	70 (34)
VPRM50	-	-	28 (28)	52 (46)
VPRM10	-	-	-	131 (100)
ORCHIDEE	-	-	-	38 (32)
5PM	-	-	-	-

Specifically, for VPRM10 and VPRM1 the e-folding time is 32 and 33 days respectively (30-34 days within 95% confidence interval for both). Confidence intervals for the e-folding time were calculated by computing the confidence intervals of the parameter in the fitted model. For ORCHIDEE best fit was 26 days (23-28 days within 95% confidence interval). In contrast, 5PM yields a significantly longer correlation time between 65-75 days (95% confidence interval) with the best fit being 70 days.

For a number of sites a large model-data bias was found. In order to assess how the result depends on individual sites where model-data residuals are more strongly biased the analysis was repeated under exclusion of sites with an annual mean of model-data flux residuals larger than 2.5μmol/m^{2}s. This threshold value is roughly half of the most deviant bias. In total 9 sites (“CH-Lae’’, “ES-ES2”, “FR-Pue”, “IT-Amp”, “IT-Cpz”, “IT-Lav”, “IT-Lec”, “IT-Ro2”, “PT-Esp”) across all model-data residuals were excluded. From these sites “CH-Lae” appears to have serious problems related to the steep terrain, where the basic assumptions made for eddy covariance flux measurements are not applicable (Göckede et al., 2008). The rest of the sites are located in the Mediterranean region, and suffer from summer drought according to the Köppen-Geiger climate classification map (Kottek et al., 2006); in those cases a large model - data bias is expected as existing models tend to have difficulties to estimate carbon fluxes for drought prone periods (Keenan et al., 2009). The model-data bias at those sites does not necessarily exceed the abovementioned threshold of 2.5μmol/m^{2}s simultaneously for each individual model, but a larger bias than the average was detected. After exclusion of those sites the temporal correlation times were found to be between 33-35 days within 95% confidence interval for 5PM with the best fit value being 34 days. The rest of the models had temporal e-folding times of 27, 29 and 24 days (1st row of Table 3.2), while the all-site correlation remains positive for lags <76, <79, <66 days for VPRM10, VPRM1 and ORCHIDEE respectively. Some weak negative correlations exist, with a minimum value of -0.06, -0.02, -0.09, -0.005 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively.

The temporal correlation of differences between VPRM10 and aircraft flux measurements could be computed for time intervals up to 36 days (Figure 3.5) corresponding to the duration of the campaign. The correlation shows an exponential decrease, and levels off after about 25 days with an e-folding correlation time of 13 days (range of 10 – 16 days within the 95% confidence interval). Whilst the general behavior is consistent with results obtained for VPRM-observation residuals for flux sites, the correlation time is two times smaller.

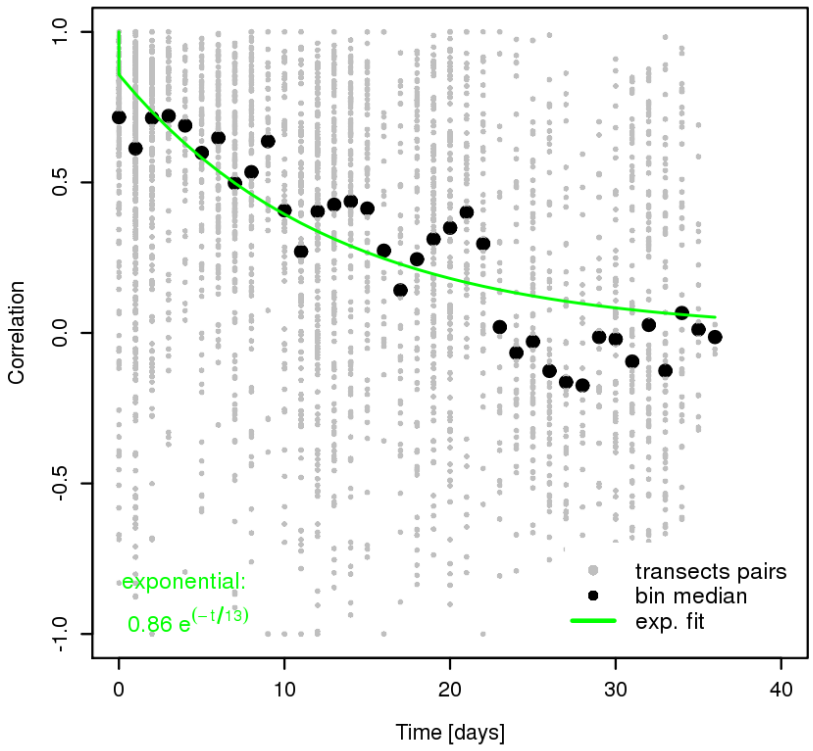

Figure 3.5 Temporal autocorrelation for VPRM10 – aircraft NEE residuals. Black dots represent individual flux transects pairs sampled at different times as function of time separation. Black circles represent daily scale binned data.

Regarding spatial error correlations, results for all models show a dependence on the distance between pairs of sites. The median correlation drops within very short distances (Figure 3.6). Fitting the simple exponentially decaying model (Eq. 3.4) to the correlation as a function of distance we find an e-folding correlation length *d* of 40, 37, 32 and 31 km with a root mean square error (RMSE) of 0.14, 0.09, 0.05 and 0.07 for VPRM10, VPRM1, ORCHIDEE and 5PM, respectively. The normalized RMSE is found to have values ranging from 0.05 to 0.084 indicating relative errors of the fit less than 9%. Spatial correlation scales are also computed for a number of different data selections (cases) in addition to the standard case shown in Figure 3.6 (case S): using only pairs with at least 150 overlapping days of non-missing data (case S*), using only pairs with identical PFT (case I), using only pairs with different PFT (case D), and using only pairs with at least 150 overlapping days for the D and I cases (cases D*, I*). The results for these cases are summarized in Figure 3.7. Also 95% confidence intervals were computed, and the spread spatial correlation was found to be markedly more critical than for the time correlations. Note that for some cases the 2.5%-ile (the lower bound of the confidence interval) hit the lower bound for correlation lengths (0 km). The e-folding correlation lengths are similar

for each of the models: this also means that no dependence on the spatial resolution was detectable. Further we examined also the spatial autocorrelation from VPRM50-data residuals with no significant difference compared to previous results.

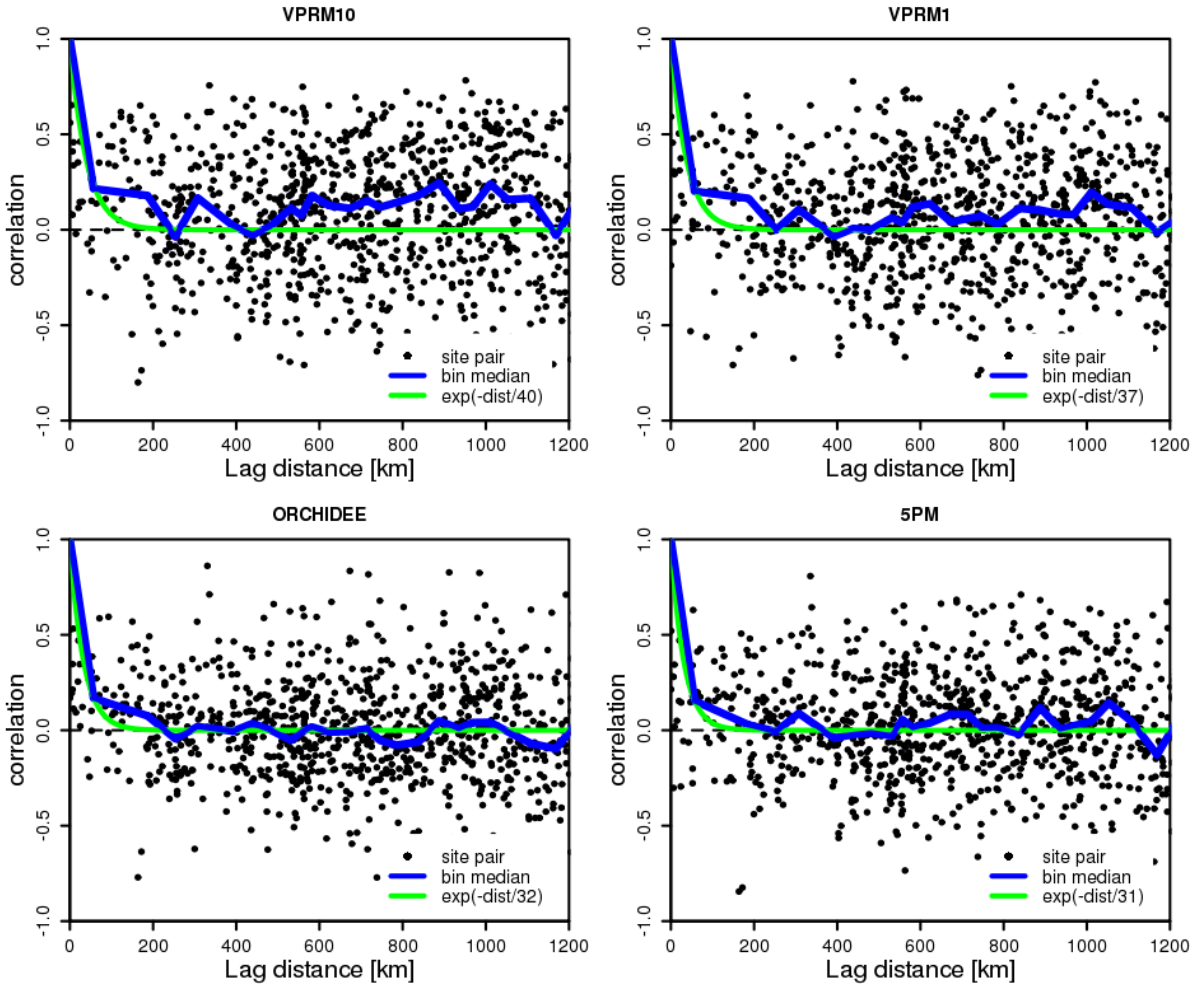

Figure 3.6 Distance correlogram for the daily net ecosystem exchange (NEE) residuals using all sites. Black dots represent the different site pairs; the blue line represents the median value of the points per 100-km bin and the green an exponential fit. Results are shown for residuals of VPRM at a resolution of 10 km (top left) and 1 km (top right), ORCHIDEE (bottom left), 5PM (bottom right).

Interestingly, if we restrict the analysis to pairs with at least 150 overlapping days between site pairs, larger correlation scales are found (case S* in Figure 3.7). Considering only pairs with different PFT (case D), consistently, all e-folding correlation lengths are found to be smaller compared to the standard case (S). This is expected to a certain degree, as model errors should be more strongly correlated between sites with similar PFTs than between sites with different PFTs. By considering only pairs within the same vegetation type (case I) we observe a significant increase of the e-folding correlation length relative to case S for VPRM at 10 and 1 km resolution to values of 432 km and 305 km, respectively. The ORCHIDEE and 5PM models show some (although not significant) increase in e-folding correlation length. Restricting again the analysis to pairs with at least 150 overlapping days for the D and I cases (D*, I*) we observe

an increase of the e-folding correlation lengths that is however significant only for VPRM at 10 and 1 km.

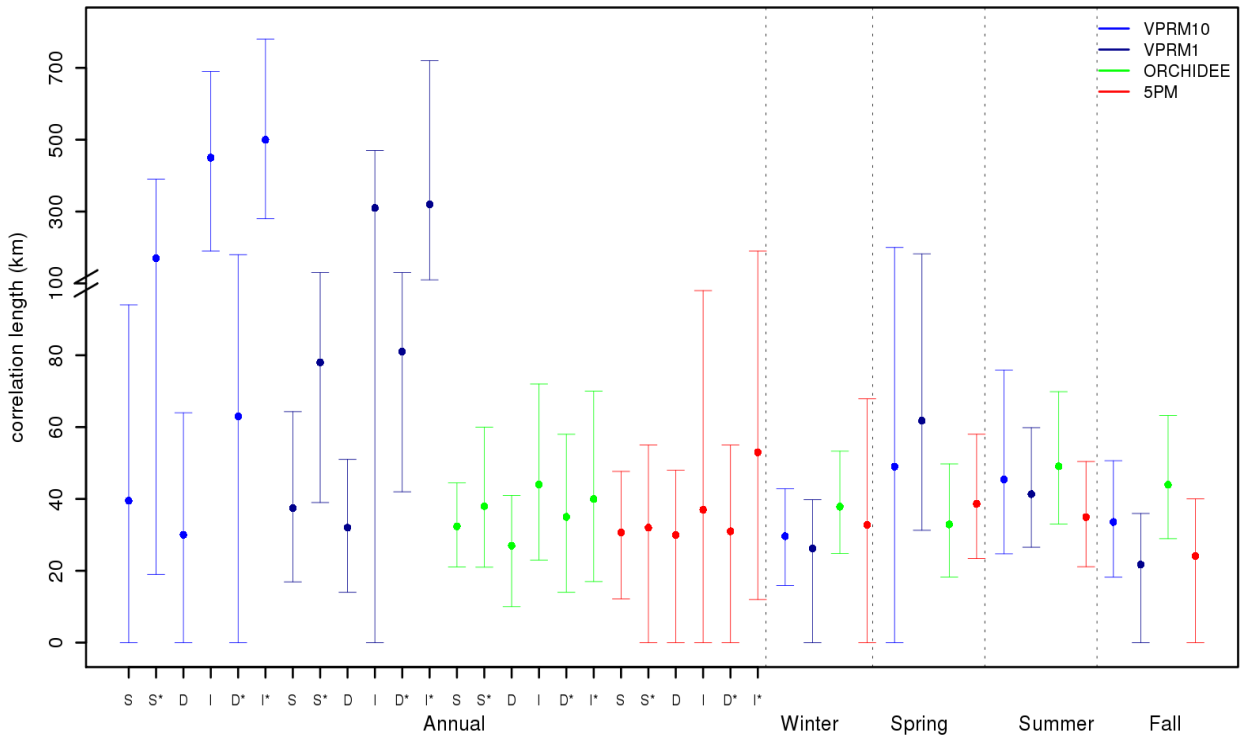

Figure 3.7 Annual and seasonal e-folding correlation length of the daily averaged model-data NEE residuals for VPRM at 10 and 1 km resolution, ORCHIDEE and 5PM. "S" refers to the standard case where all pairs were used, "D" refers to the case where only pairs with different vegetation types were used, "I" denotes the case in which only pairs with identical vegetation type were considered, and "*" denotes that in addition 150 days of common non-missing data are required for each pair of sites. The dot represents the best-fit value when fitting the exponential model. The upper and the lower edge of the error bars show the 2.5 and 97.5 percentiles of the length value. Note the scale change in the y-axis at 100 km.

Seasonal dependence of the e-folding correlation lengths for at least 20 overlapping days per season and for all site-pairs is also shown in Figure 3.7. VPRM showed somewhat longer correlation lengths during spring and summer, ORCHIDEE had the largest lengths occurring during summer and fall, and 5PM e-folding correlation lengths show slightly enhanced values during spring and summer. However, none of these seasonal differences are significant with respect to the 95% confidence interval.

The spatial error correlation between VPRM10 model and aircraft fluxes measured during May-June along continuous transects at forest and agriculture land use (Figure 3.8) shows an exponential decay up to the maximum distance that was encompassed during flights (i.e. 70 km). Of note is that only two measurements were available at 60 km distance and none for larger distances making it difficult to identify where the asymptote lies. Nevertheless fitting the decay model (Eq. 3.4) leads to d = 35km (26 – 46 km within the 95% confidence interval), which is in good agreement with the spatial correlation scale derived for VPRM10 using flux sites during both spring and summer (Figure 3.7).

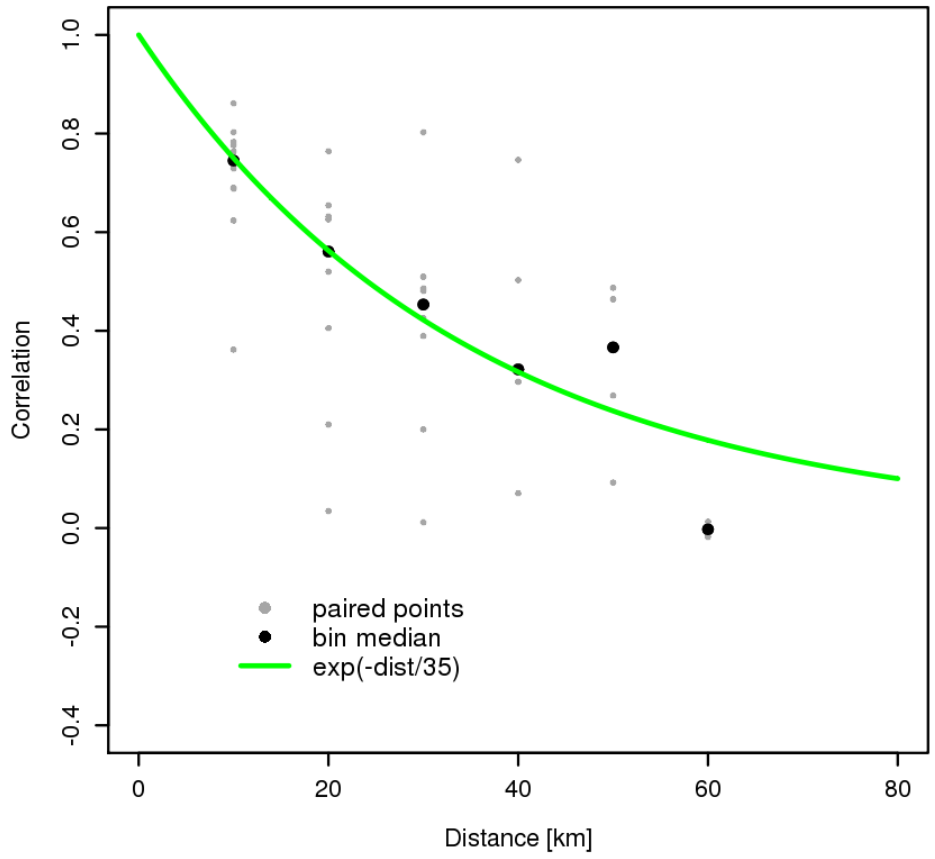

Figure 3.8 Distance correlogram between VPRM10 and aircraft NEE measurements. Black dots represent the different aircraft grid points pairs; black circles represent 10 km scale binned data.

3.3.2 Model-model comparison

We investigate the model-model error structure of NEE estimates by replacing the observed fluxes which were used as reference, with simulated fluxes from all the biosphere models. Note that for consistency with the model-data analysis, the simulated fluxes contained the same gaps as the observed flux time series. The e-folding correlation time is found to be slightly larger compared to the model-data correlation times, for most of the cases. An exception is the 5PM-VPRM10 pair which produced remarkably larger correlation time (Table 3.2). Specifically, VPRM50-ORCHIDEE and VPRM10-5PM residuals show correlation times of 28 days (range between 24-32 days within 95% confidence interval) and 131 (range between 128-137 days within 95% confidence interval), respectively. Significantly different e-folding correlation times are found for VPRM50-5PM compared to VPRM10-5PM with correlation times of 52 days (range between 49-56 days within 95% confidence interval).Repeating the analysis excluding sites with residual bias larger than 2.5μmol/m^2s, correlation times of 28 and 100 days for VPRM50-ORCHIDEE and VPRM10-5PM are found, respectively. If we use ORCHIDEE-5PM pair the e-folding correlation time found to be 38 days (range between 35-41 days within 95% confidence interval).

Although the e-folding correlation times show but minor differences compared to the model-data residuals, this is not the case for the spatial correlation lengths (Figure 3.9). The standard case

(S) was applied for the annual analysis, with no minimum number of days with overlapping non-missing data for each site within the pairs. Taking VPRM50 as reference, much larger e-folding correlation lengths of 371 km with a range of 286-462 km within 95% confidence interval yielded for VPRM50-ORCHIDEE comparisons, and 1066 km for VPRM50-5PM were found. However VPRM10-5PM analysis which is also considered appropriate in terms of the spatial resolution compatibility contrary to the VPRM50-5PM pair, is in good agreement with VPRM50-ORCHIDEE spatial scale (230-440 km range within 95% confidence interval with the best fit being 335 km). With ORCHIDEE as reference the e-folding correlation length for the ORCHIDEE-5PM comparison is 276 km with a range of 183-360 km within 95% confidence interval. However the later correlation length might be affected by the different spatial resolution as the difference between VPRM10 and VPRM50 against 5PM suggests. Seasonal e-folding correlation lengths, using a minimum of 20 days overlap in the site-pairs per season (Figure 3.9), are also significantly larger compared with those from the model-data analysis.

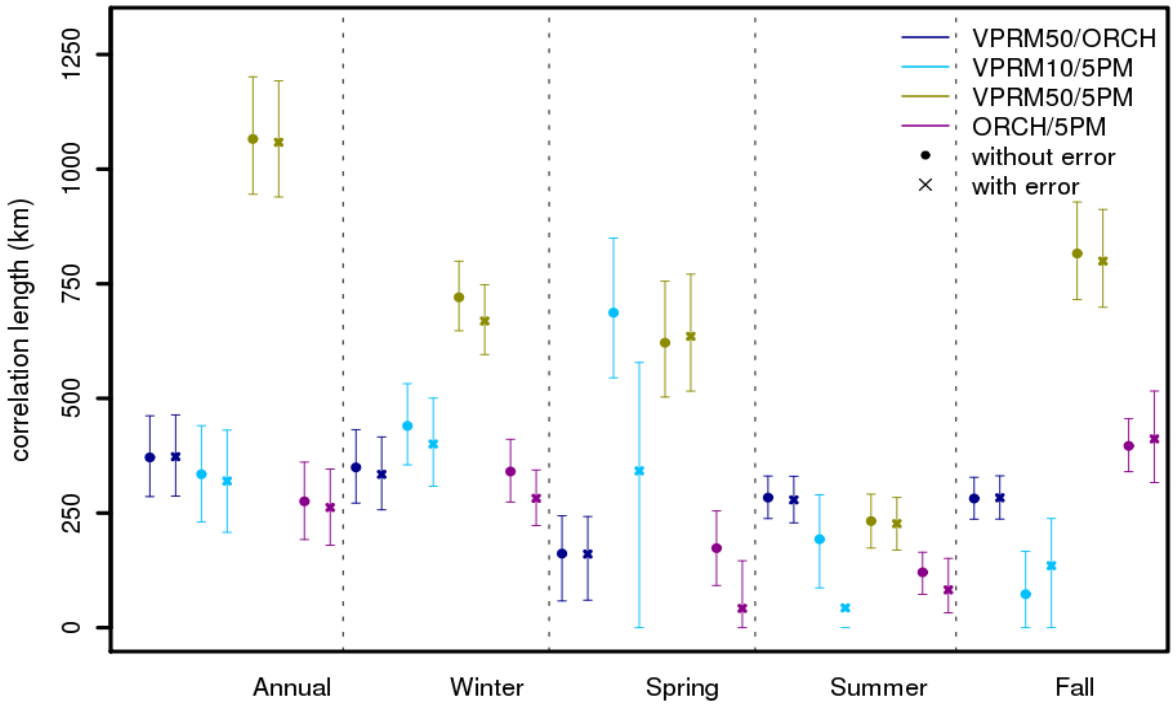

Figure 3.9 Annual and seasonal e-folding correlation length for an ensemble of daily averaged NEE differences between two models without (filled circle) and with random measurement errors added to the modeled fluxes used as reference (crosses). The symbols represents the best fit value when fitting the exponential model, and the upper and lower edge of the error bars show the 2.5 and 97.5 percentiles of the correlation length. The first acronym at the legend represents the model used as reference and the second the model which was compared with. Note that for the VPRM10/VPRM1 case during spring (with and without random error), the 97.5 percentile of the length value exceeds the y-axis and has a value of 1073, 1626 km respectively.

When we add the random measurement error to the modeled fluxes used as reference (crosses in Figure 3.9), we observe only slight changes in the annual e-folding correlation lengths, without a clear pattern. The correlation lengths show a random increase or decrease but limited up to 6%. Interestingly, the seasonal e-folding correlation lengths for most of the cases show a more clear

decrease. For example, the correlation length of the VPRM10-5PM residuals during winter, decreases by 22% or even more for spring season. Despite this decrease, the e-folding seasonal correlation lengths remain significantly larger in comparison to those from the model-data analysis. Overall, all models when used as reference show the same behavior with large e-folding correlation lengths that mostly decrease slightly when the random measurement error is included. Although the random measurement error was added as "missing part" to the modeled fluxes to better mimic actual flux observations, it did not lead to correlation lengths similar to those from the model-data residual analysis. To investigate if a larger random measurement error could cause spatial correlation scales in model-model differences, we repeated the analysis with artificially increased random measurement error (multiplying with a factor between 1 and 15). Only for very large random measurement errors did the model-model e-folding correlation lengths start coinciding with those of the model-data residuals (Figure 3.10).

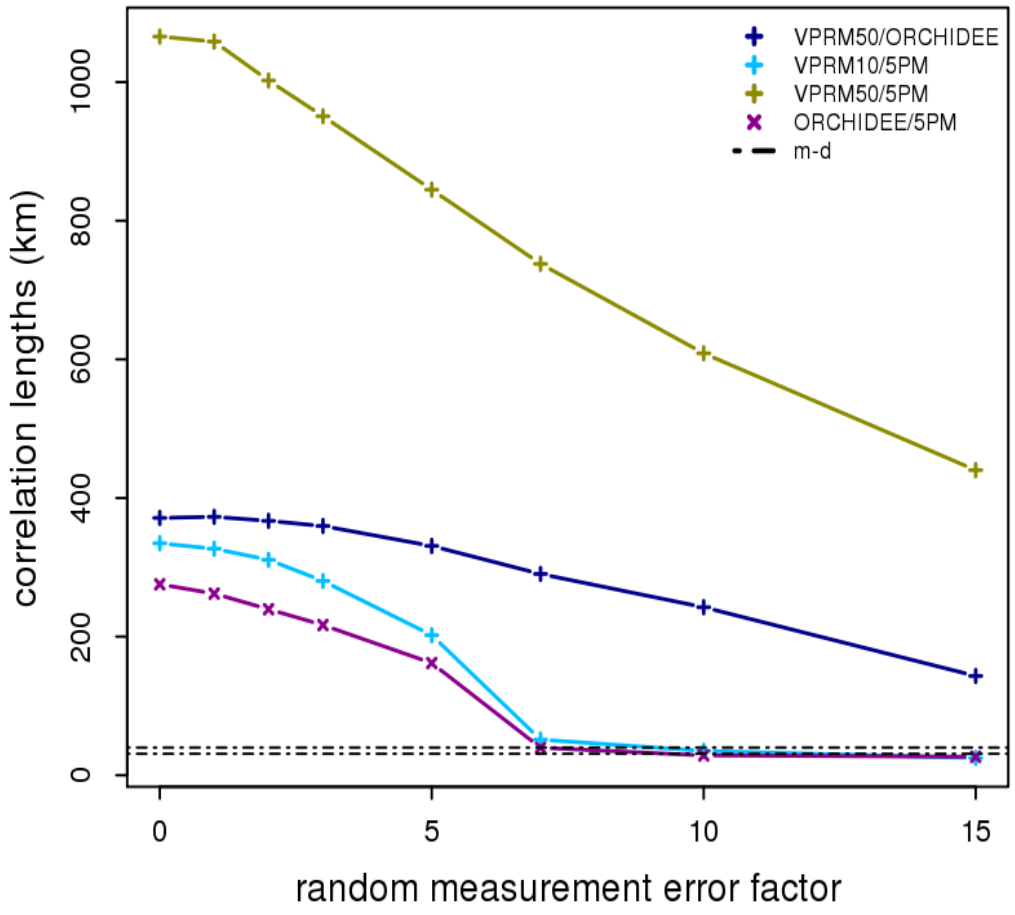

Figure 3.10 Annual e-folding correlation lengths as a function of the factor used for scaling the random measurement error, for all model-model combinations. The black dot-dash lines reveal the range of the spatial correlation lengths generated from the model-data comparisons.

3.4 Discussion and conclusions

We analyzed the error structure of a-priori NEE uncertainties derived from a multi-model – data comparison by comparing fluxes simulated by three different vegetation models to daily

averages of observed fluxes from 53 sites across Europe, categorized into 7 land cover classes. The different models showed comparable performance with respect to reproducing the observed fluxes; we found mostly insignificant differences in the mean of the residuals (bias) and in the variance. Site-specific correlations between simulated and observed fluxes are significantly higher than overall correlations for all models, which suggest that the models struggle with reproducing observed spatial flux differences between sites. Furthermore, the site-specific correlations reveal a large spread even within the same vegetation class, especially for crops (Figure 3.2). This is likely due to the fact that none of the models uses a specific crop model that differentiates between the different crop types and their phenology. The models using remotely sensed vegetation indices (VPRM and 5PM) better capture the phenology; ORCHIDEE is the only model that differentiates between C_3 and C_4 plants, but shows the largest spread in correlation for the crop. Differences in correlations between the different vegetation types were identified for all the biosphere models, however it must be noted that the number of sites per vegetation type is less than 10 except for crop and evergreen forests.

Model-data flux residual correlations were investigated to give insights regarding prior error temporal scales which can be adopted by atmospheric inversion systems. Whilst fluxes from ORCHIDEE model are at much coarser resolution compared to the representative area from the flux measurements, VPRM1 fluxes (1 km resolution and only the meteorology at 25 km) are considered appropriate for the comparisons. Despite the scale mismatch results are in good agreement across all model-data pairs.

Exponentially decaying correlation models are a dominant technique among atmospheric inverse studies to represent temporal and spatial flux autocorrelations (Rödenbeck et al., 2009, Broquet et al., 2011, Broquet et al., 2013). However, regarding the temporal error structure we need to note the weakness of this model to capture the slightly negative values at 2-10 months lags and, more importantly, the increase in correlations for lag times larger than about 10 months. Error correlations were parameterized differently by Chevalier et al., (2012) where the prior error was investigated without implementing it to atmospheric inversions. Polynomial and hyperbolic equations were used to fit temporal and spatial correlations respectively. Nevertheless, we use here e-folding lengths not only for their simplicity in describing the temporal correlation structure with a single number, but also because this error model ensures a positive definite covariance matrix (as required for a covariance). This is crucial for atmospheric inversions as otherwise negative, spatially and temporally integrated uncertainties may be introduced. In addition it can keep the computational costs low; this is because the hyperbolic equation has significant contributions from larger distances: for the case of the VPRM1 model, at 200 km distance the correlation according to Chevallier et al., hyperbolic equation is 0.16, compared to 0.004 for the exponential model. As a consequence, more none-zero elements are introduced to the covariance matrix, which increases computational costs in the inversion systems. Using the same hyberbolic equation for the spatial correlation, d-values of 73, 39, 12 and 20 km were found with a RMSE of 0.11, 0.07, 0.05, 0.07 for VPRM10, VPRM1, ORCHIDEE and 5PM respectively. A similar RMSE was found when using the exponential (0.14, 0.09, 0.05 and 0.07), indicating similar performance of both approaches with respect to fitting the spatial correlation.

Autocorrelation times were found to be in line with findings of Chevallier et al., (2012). The model-data residuals were found to have an e-folding time of 32 and 26 days for VPRM and

ORCHIDEE respectively, and 70 days for 5PM. This significant difference appears to have a strong dependence on the set of sites used in the analysis. Excluding nine sites with large residual bias, the autocorrelation time from the 5PM-data residuals drastically decreased and became coherent with the times of the other biosphere models. The all-models and all-sites autocorrelation time was found to be 39 days, which reduces to 30 days (28-31 days within 95% confidence interval), when excluding the sites with large residual bias, coherent with the single model times. From the model-model residual correlation analysis, the correlation time appear to be consistent with the above-mentioned results, and lies between 28 and 46 days for most of the ensemble members. However model-model pairs consisting of the VPRM and 5PM models produced larger times up to 131 days; omitting sites with large residual biases this is reduced to 100 days (99-105 days within 95% confidence interval). This finding could be attributed to the fact that despite the conceptual difference between those models, they do have some common properties. Both models were optimized against eddy covariance data although for different years (2005 and 2007 respectively), while no eddy covariance data were used for the optimization of ORCHIDEE. In addition, VPRM and 5PM both use data acquired from MODIS, although they estimate photosynthetic fluxes by using different indices of reflectance data. Summarizing the temporal correlation structure, it appears reasonable to a) use same error correlation in atmospheric inversions regardless which biospheric model is used as prior, b) use an autocorrelation length of around 30 days.

Only weak spatial correlations for model-data residuals were found, comparable to those identified by Chevallier et al. (2012) limited to short lengths up to 40 km without any significant difference between the biospheric models (31 - 40 km). Hilton et al. (2012) estimated spatial correlation lengths of around 400km. However we note that significant differences exist between this study and Hilton et al. (2012) regarding the methods that were used and the landscape heterogeneity of the domain of interest. With respect to the first aspect the time resolution is much coarser (seasonal averaged flux residuals) compared to the daily averaged residuals used here. Furthermore spatial bins of 300 km were used for the autocorrelation analysis, which is far larger than the approximate bin width of 100 km that were used in our study. Regarding the second aspect North America has a more homogenous landscape compared to the European domain. The scales for each ecosystem type (e.g. forests, agricultural land etc.) are drastically larger than those in Europe as can be seen from MODIS retrievals (Friedl et al., 2002).

Although the estimated spatial scales are shorter than the spatial resolution that we are solving for (100 km bins), the autocorrelation analysis of aircraft measurements made during CERES supports the short scale correlations. These measurements have the advantage of providing continuous spatial flux transects along specific tracks that were sampled routinely (in this case over period of 36 days at various times of the day), thus resolving flux spatial variability also at small scales, where pairs of eddy covariance sites may not be sufficiently close. On the other hand, aircraft surveys are necessarily sporadic in time. Of note is that the eddy covariance observation error has no significant impact on the error structure, as the addition of an observation error to the analysis of model-model differences had only minor influence on the error structure. We note that the current analysis focuses to daily time scale and therefore the error statistics with respect to the estimated spatial and temporal e-folding correlation lengths are valid for such scales.

Model-data residual e-folding correlation lengths show a clear difference, between the cases where pairs only with different (D) or identical (I) PFT were considered, with the latter resulting in longer correlation lengths, but only identified for the VPRM model at both resolutions. The "D" case has slightly shorter lengths for all models than the standard case (S). One could argue that as VPRM uses PFT specific parameters that were optimized against 2005 observations, the resulting PFT specific bias could lead to longer spatial correlations. However ORCHIDEE and 5PM also show comparable biases (Figure 3.3), but long correlation scales were not found. Moreover we repeated the spatial analysis after subtracting the PFT specific bias from the fluxes, and the resulting correlation lengths showed no significant change. The impact of data gaps was also investigated by setting a threshold value of overlapping observations between site pairs. Setting this to 150 days results in an increase for the "S" case up to 60 km, but only for the VPRM model. For the "D" and "I" cases when setting the same threshold value (D^* and I^*) we only found an insignificant increase, indicating that data gaps are hardly affecting the "D" and "I" cases. These findings suggest that high-resolution diagnostic models might be able to highlight the increase of the spatial correlation length between identical PFTs vs. different PFTs. Note that the Chevallier et al., (2012) study concluded that assigning vegetation type specific spatial correlations is not justified, based on comparisons of eddy covariance observations with ORCHIDEE simulated fluxes. The current study could not further investigate this dependence, as the number of pairs within a distance bin is not large enough for statistical analyses, when using only sites within the same PFT. With respect to the seasonal analysis, spatial correlations are at the same range among all models and seasons. Although in some cases (VPRM10 and VPRM1 spring) the scales are larger, they suffer from large uncertainties. Hence, implementing distinct and seasonally dependent spatial correlation lengths in inversion systems cannot be justified.

The analysis of model-model differences did not reproduce the same spatial scales as those from the model-data differences, but instead spatial e-folding correlation lengths were found to be dramatically larger. Adding a random measurement error to the modeled fluxes used as reference slightly reduced the spatial correlation lengths to values ranging from 278 to 1058 km. Even when largely inflating the measurement error, the resulting spatial correlation lengths (Figure 3.10) still do not approach those derived from model-data residuals. Only when the measurement error is scaled up by a factor of 8 or larger (which is quite unrealistic as this corresponds to a mean error of 1.46 μmol m^{-2} s^{-1} or larger, which is comparable to the model-data mismatch where a standard deviation of around 2.5 μmol m^{-2} s^{-1} was found), the e-folding correlation lengths are consistent with those based on model-data differences. Whilst the EC observations are sensitive to a footprint area of about 1 km^2, the model resolution is too coarse to capture variations at such a small scale. This local uncorrelated error has not been taken into account by the analysis of model-data residuals as the error model could not be fitted with a nugget term included, favoring therefore smaller correlation scales. The analysis of differences between two coarser models does not involve such a small scale component, thus resulting in larger correlation scales. This would suggest that for inversion studies targeting scales much larger than the eddy covariance footprint scale, the statistical properties of the prior error should be derived from the model-model comparisons.

The large e-folding correlation lengths yielded from this model-model residual analysis suggest that the models are more similar to each other than to the observed terrestrial fluxes, at least on

spatial scales up to a few hundred kilometers regardless of their conceptual differences. This might be expected to some extent due to elements that the models share. Respiration and photosynthetic fluxes are strongly driven by temperature and downward radiation, respectively, and those meteorological fields have significant commonalities between the different models. VPRM and 5PM both use temperature and radiation from ECMWF analysis and short-term forecasts. Also the WFDEI temperature and radiation fields used in ORCHIDEE are basically from the ERA-Interim reanalysis, which also involves the integrated forecasting system (IFS) used at ECMWF (Dee et al., 2011). Regarding the vegetation classification all models are site specific and therefore are using the same PFT for each corresponding grid-cell. Photosynthetic fluxes are derived with the use of MODIS indices in VPRM (EVI and LSWI) and in 5PM (LAI and albedo).

Using full flux fields from the model ensemble (rather than fluxes at specific locations with observation sites only) to assess spatial correlations in model-model differences is not expected to give significantly different results, as the sites are representative for quite a range of geographic locations and vegetation types within the domain investigated here.

The current study intended to provide insight on the error structure that can be used for atmospheric inversions. Typically, inversion systems have a pixel size ranging from 10 to 100 km for regional and continental inversions, and as large as several degrees (hundreds of km) for global inversions. If a higher resolution system assumes such small-scale correlations (as those found in the current analysis), in the covariance matrix, of note is that this leads to very small prior uncertainties when aggregating over large areas and over longer time periods. To aggregate the uncertainty to large temporal and spatial scales, we used the following equation (after Rodgers, 2000):

$$Ua = u \times Q_c \times u^T \qquad \textbf{3.7}$$

Where "×" denotes matrix multiplication, Q_c is the prior error covariance matrix and u a scalar operator that aggregates the full covariance to the target quantity (e.g. domain-wide and full year). For example, with a 30 km spatial and a 40 day temporal correlation scale, annually and domain-wide (Figure 3.1) aggregated uncertainties are around 0.06 GtC. This is about a factor ten smaller than uncertainties typically used e.g. in the Jena inversion system (Rödenbeck et al., 2005). This value is also 8 times smaller when comparing it to the variance of the signal between 11 global inversions reported in Peylin et al., (2013) which was found to be 0.45 GtC/y, proving that the aggregated uncertainties are unrealistically small. In addition, the aggregated uncertainties using the VPRM10-ORCHIDEE error structure (32 days and 320 km temporal and spatial correlation scales) are found to be 0.46 GtC/y which is also much smaller than the difference between VPRM10 (NEE= - 1.45 GtC/y) and ORCHIDEE (NEE= - 0.2 GtC/y), when aggregated over the domain shown in Figure 3.1. Although this analysis does capture the dominating spatiotemporal correlation scale in the error structure, it fails in terms of the error budget, suggesting that also other parts of the error structure are important as well. Therefore additional degrees of freedom (e.g. for a large-scale bias) need to be introduced in the inversion systems to fully describe the error structure.

Whilst temporal scales found from this study have already been used in inversion studies, this is not the case to our best knowledge for the short spatial scales. The impact of the prior error structure derived from this analysis, on posterior flux estimates and uncertainties will be assessed in a subsequent paper. For that purpose, findings from this study are currently implemented in three different regional inversion systems aiming to focus on network design for the ICOS atmospheric network.

Chapter 4

Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part1: Methodology and system evaluation

Abstract

Atmospheric inversions are widely used in the optimization of surface carbon fluxes at regional scale using information from atmospheric CO_2 dry mole fractions. In many studies the prior flux uncertainty applied to the inversion schemes does not reflect directly the true flux uncertainties but it is used in such a way to regularize the inverse problem. Here, we aim to implement an inversion scheme using the Jena inversion system and applying a prior flux error structure derived from a model – data residual analysis using high spatial and temporal resolution over a full year period in the European domain. We analyzed the performance of the inversion system with a synthetic experiment, where the flux constraint is derived following the same residual analysis but applied to the model-model mismatch. The synthetic study showed a quite good agreement between posterior and "true" fluxes at European/Country and annual/monthly scales. Posterior monthly and country aggregated fluxes improved their correlation coefficient with the "known truth" by 7% compared to the prior estimates when compared to the reference, with a mean correlation of 0.92. Respectively, the ratio of the standard deviation between posterior/reference and prior/reference was also reduced by 33% with a mean value of 1.15. We identified temporal and spatial scales where the inversion system maximizes the derived information; monthly temporal scales at around 200 km spatial resolution seem to maximize the information gain.

4.1 Introduction

The continuous rise of the abundance of greenhouse gases in the atmosphere, especially due to fossil fuel combustion, alerted the scientific community to systematically monitor these emissions. The challenge is not limited only to revealing the spatial distribution of CO_2 sources and sinks on continental scales, but also to accurately quantifying CO_2 emissions and their uncertainties at country scales. In situ atmospheric measurements of the atmospheric CO_2 variability combined with inverse atmospheric models are used as an independent method to provide "top down" flux estimates for comparison with estimates from "bottom up" methods. The latter use local observations (e.g. eddy covariance), and combine these with ancillary data, e.g. soil maps, satellite data, and terrestrial ecosystem models in order to spatially scale up local flux estimates to larger regions (Jung et al., 2009). Both approaches act complementary, for optimal comprehension of carbon sources and sinks in a "multiple constraint" (Schulze et al., 2010) approach and emission inventories assessment. As these inventories are used to deduce national emission estimates, in compliance with the Kyoto protocol requirements, accuracy is essential.

An atmospheric inverse modeling system provides the link from atmospheric concentrations to surface fluxes. However, the limited number of observations available for solving the system for quite a number of unknowns (spatially and temporally resolved fluxes) makes the inverse problem strongly under-determined. To solve the inverse problem the system incorporates Bayes' theorem and uses a-priori knowledge, provided by e.g. biosphere models and emission inventories accompanied by corresponding uncertainty estimates. Then, the system optimizes the a-priori fluxes by minimizing the difference between model predictions and observed concentrations. For the current study only the biospheric fluxes were optimized, and emissions from fossil fuel combustion are assumed to be known much better, as it is the case in almost all published regional inversion studies. Inversion systems have been extensively used to derive spatiotemporal flux patterns at global (e.g. Enting et al., 1995; Kaminski et al., 1999a; Gurney et al., 2003; Mueller et al., 2008), and regional scale (e.g. Gerbig et al., 2003a; Peylin et al., 2005; Lauvaux et al., 2012; Broquet et al., 2013).

The challenge in regional inversions is to reconstruct at high resolution the spatiotemporal flux patterns, usually of the net ecosystem exchange (NEE). For that purpose currently deployed global or regional inverse modeling schemes use different state spaces (i.e. the set of variables to be optimized through the inversion process). Peters et al. (2007) split the domain of interest into regions according to ecosystem type. Subsequently fluxes are optimized by using linear multiplication factors to scale NEE for each week and each region. The pitfall of this system is that a zero prior flux has no chance to be optimized and remains zero. Zupanski et al. (2007) divided the NEE into two components, i.e. the gross photosynthetic production (GPP) and ecosystem respiration (R). Then multiplicative factors for the gross fluxes were derived on the grid scale, under the assumption of being constant in time. A step further made by Lokupitiya et al. (2008) used the same approach but with an 8-week time window allowing for temporal variations for the multiplicative factors. A different approach introducing the carbon cycle data assimilation system (CCDAS) was implemented by Rayner et al. (2005) and Kaminski et al. (2012) by constraining global parameters within a biosphere model able to control surface-atmosphere exchange fluxes, against observed atmospheric CO_2 mole fractions, instead of the fluxes themselves. Lauvaux et al. (2012) used a Bayesian approach based on matrix inversion,

separately optimizing day and night time fluxes at a weekly time scale for a limited simulation period and domain. An attempt to assess which of these approaches better reproduces NEE was made by Tolk et al. (2011). This study investigated the impact of different inversion approaches via a synthetic experiment utilizing an ensemble Kalman filter technique and the same transport model for all cases. They found that inversions which separately optimize gross fluxes within a pixel inversion concept perform better on reconstructing the NEE, although they fail to obtain the gross fluxes. Taking into consideration these findings we also choose the pixel based inversions but optimizing the net biogenic fluxes as we are mainly interested in the total carbon flux budget.

Introducing proper prior flux uncertainties is crucial for meaningful posterior estimates, as these uncertainties weight the prior knowledge between different locations and times, as well as with respect to the data constraint. The uncertainties have the form of a covariance matrix and can be categorized in uncertainties of the prior fluxes, and uncertainties of the observational constraint, which includes measurement and transport model uncertainties. While the observational constraint may be more easily defined with the main diagonal of the covariance matrix representing the uncertainty of the observations and the model at a specific time and location, our knowledge for the prior uncertainty is limited. Early inversions assumed fully uncorrelated flux uncertainties (Kaminski et al., 1999b), while spatial and temporal correlations were used later by Rödenbeck et al. (2003), who investigated the autocorrelation of monthly CO_2 fluxes calculated by a set of terrestrial and ocean models. In Rödenbeck (2005), spatial correlations for land fluxes were assigned to a state space of 4° latitude x 5° longitude resolution. Slightly different correlation length scales were considered for the meridional and zonal direction, assuming that the climate zone of the later varies less than of the former. Flux correlations on land were determined by assuming an exponential pulse response function with a length of 1275 km. This leads to correlations with approximately twice the correlation length. Typically the spatial correlations are considered more as a tool to regularize the inverse problem, rather than an uncertainty feature. Schuh et al. (2010) obtained correlation lengths from Rödenbeck et al. (2003) but with a much higher state space resolution of 200 km. Lauvaux et al. (2008) neglected the spatial correlations to enlarge the impact of the data. Carouge et al. (2010a) inferred spatial and temporal correlation lengths based on the agreement between posterior and "true" fluxes in the framework of a synthetic experiment, where the "truth" is known. A different approach was used in Peters et al. (2007) study where they interpret the length scale from a climatological and ecological perspective, and use it to spread information within regions, which the network is incapable to constrain. Ad-hoc solutions have also been used, assuming that daily fluxes have smaller correlation lengths than monthly fluxes which are used by other studies (Peylin et al. 2005). More specifically Peylin et al. (2005) assumed 500 km for daily temporal resolution compared to the much larger correlation lengths used by Rödenbeck for monthly flux resolution. Michalak et al. (2004) implemented a geostatistical approach to describe the prior error structure. Specifically the prior error covariance describes at which degree deviations of the surface fluxes from their mean behavior at two different locations or times are expected to be correlated as a function of the distance in space or in time. They simultaneously estimate posterior fluxes as well as parameters controlling the model-data mismatch uncertainty and the prior flux uncertainty, including spatial and temporal correlation lengths. Although this approach may be considered as an objective way to infer spatial and temporal correlation lengths, it forces the error covariance to be statistically consistent with the atmospheric data from the few regions where station-to-station distances are small enough to be comparable to the correlation length

scales. Eddy Covariance stations (EC) can provide a more direct method to infer spatial and temporal flux correlations. Chevallier et al. (2006) and Chevallier et al. (2012) introduced autocorrelation analysis of the residual between fluxes simulated by biosphere models or measured by EC to infer spatial and temporal error correlations. The derived error statistics were implemented in a regional CO_2 inversion by Broquet et al. (2013).

Daily NEE flux residuals from model - data comparisons showed temporal correlations up to 30 days but very short spatial correlations up to 40 km (Kountouris et al. 2015). In such a case the a-priori integrated uncertainty over time and space, e.g. annually and EU wide domain integrated, according to the error propagation will be exceptionally small. For example a variance of 1.82 $\mu mole.m^{-2}.s^{-1}$ (from model – data differences) combined with the abovementioned correlation scales yields an uncertainty of 0.12 GtC y^{-1} for the total flux over Europe. This value is significantly smaller than the assumed uncertainty which is typically used by the inversion systems. For comparison we refer to studies from Rivier et al. (2010) and Peylin et al. (2005) (for a slightly larger domain than ours) where an a priori uncertainty of approximately 1.4 GtC y^{-1} and 1 GtC y^{-1} respectively was used. Further, Peylin et al. (2013) found that the variance of the posterior NEE fluxes for the European domain among 11 global inversions is also 3 to 4 times larger (0.45 GtC y^{-1}). Although is not yet entirely clear what would be the "correct" value for the prior uncertainty, it seems that in our study it should be increased not only to give enough flexibility to the system to adjust but also to be at least comparable with other posterior uncertainty estimates. A typical method is to inflate the spatiotemporal component by scaling accordingly the prior error covariance. In a study by Lauvaux et al. (2012) two correlation lengths were used at 300 and 50 km, and for the shorter scale the uncertainty was inflated by increasing the RMS of the prior error covariance. The model - data analysis (Kountouris et al. 2015) does neither justify the use of large correlation scales nor largely inflated variances which exceed the model-data flux mismatches, however it is consistent with an additional overall bias error which can not be captured from the estimated spatiotemporal error structure. Hence an appropriate approach would be to introduce two adjustable terms into the inversion system. One term to reflect the data-derived error structure without error inflation (prior error covariance matrix which describes the spatiotemporal component) and one term to represent a bias component. To the best of our knowledge such an approach has not yet been used in inversion systems.

This study primarily aims to use the information extracted from the model-EC data residuals (spatiotemporal error structure) to define a data-driven error covariance rather than simply assuming one, adopting a conservative one or an expert knowledge solution. For that, we implement our previous methodology and findings regarding the prior uncertainty to atmospheric inversions following Kountouris et al. (2015). As explained above, we implement two uncertainty terms; the first one to reflect the true spatiotemporal error structure and the second term referred to a bias term. We use the Jena inversion system (Rödenbeck, 2005; Rödenbeck et al., 2009) for the regional scale consisting of a fully coupled system as described in Trusilova et al. (2010), between the global three-dimensional atmospheric tracer transport model TM3 (Heimann and Körner, 2003) and the regional stochastic Lagrangian transport model STILT (Lin et al., 2003). This scheme allows retrieving surface fluxes at much finer resolution (0.25^o) compared to global models. The first part of this study details the methodology of the prior error implementation, and evaluates the system's performance through a synthetic data experiment. The system evaluation is an extension of Trusilova et al. (2010) where the evaluation was limited

to the observation space only. We extend that to the flux space by comparing flux retrievals at various spatial and temporal scales against synthetic "true" fluxes. Station locations and observation times (including gaps) were created as in the real observation time series presented in the second part of this study (Kountouris et al., 2016). That way we can use the synthetic experiment to evaluate to what extent we can trust the results, if a real-data inversion is performed. In the second part of this study (Kountouris et al., 2016) the regional inversion system is applied to real observations of atmospheric CO_2 mole fractions from a network of 16 stations.

This chapter is structured as follows. In Section 4.2 we present the inversion scheme and introduce the settings of the atmospheric inversions. In Section 4.3 we present the results from a synthetic inversion experiment aimed to assess the prior error setup, considering it as a step towards atmospheric inversions using real atmospheric data with an objective, state of the art prior error formulation. Discussion and conclusions are following in Section 4.4.

4.2 Methods

4.2.1 Inversion scheme

The Jena Inversion System (Rödenbeck 2005; Rödenbeck et al., 2009) was used for the current study. The scheme is based on the Bayesian inference and uses two transport models, the TM3 model (Heimann and Körner, 2003) for global, and the STILT model (Lin et al., 2003) for regional simulations. The advantage of the system is that it combines a global transport model with a regional one without the need of a direct coupling along the boundaries. The global is used to calculate fluxes from the far field (outside of the regional domain of interest), and subsequently this information can be used to provide lateral boundary information for the regional model. Primary input of the system is the observed mixing ratios c_{meas}. This vector contains all measured mixing ratios at different times and locations. The modeled mixing ratios c_{mod} given from a temporally and spatially varying discretized flux field f are computed from an atmospheric transport model and can be formally expressed as

$$c_{\mathrm{mod}} = Af + c_{ini} \tag{4.1}$$

where c_{ini} is the initial concentration and A the transport matrix which maps the flux space to the observation space. For the regional domain the transport matrix A has been pre-computed by the STILT transport model. The system calculates the modeled concentrations when and where a measurement exists in the c_{meas} vector.

In the following, we briefly describe the inverse modeling approach. For more details the reader is referred to Rödenbeck (2005).

In grid-based atmospheric inversions the number of unknowns (spatially and temporally resolved fluxes) is larger than the number of measurements (hourly dry mole fractions at different sites), making the inverse problem ill-posed. In the Bayesian concept this can be remedied by adding a-priori information. This information can be written as

$$f = f_{fix} + F \cdot p$$

4.2

where f_{fix} is the a-priori expectation value of the flux, matrix F contains all the a-priori information about flux uncertainties and correlations (implicitly defining the covariance matrix) and p is a vector representing the adjustable parameters. The parameters p are uncorrelated with zero mean and unit variance. This flux model represents just a different way to define the a-priori probability distribution of the fluxes, than the traditional way where the a-priori error covariance matrix is explicitly specified. The cost function describing the observational constrain is expressed as

$$J_c = \frac{1}{2}(c_{meas} - c_{\mathrm{mod}})^T \cdot Q_c^{-1} \cdot (c_{meas} - c_{\mathrm{mod}})$$

4.3

where Q_c is the observation error covariance matrix. This diagonal matrix weights the mixing ratio values considering measurement uncertainty, location-dependent model uncertainty and a data density weighting. The latter ensures that the higher amount of data from continuous measurements compared to the data from flask measurements would not lead to a considerably stronger impact of these corresponding sites (Rödenbeck, 2005). This can also be formally interpreted as a temporal correlation scale which ensures that the model-data-mismatch error is not independent within a week, corresponding roughly to time scales of synoptic weather patterns.

The inversion system seeks to minimize the following cost function that combines the observational (Eq. 4.3) and the prior flux constrain

$$J = J_c + \frac{1}{2} \cdot p^T \cdot p$$

4.4

The minimization of the cost function is done iteratively with respect to the parameters p by using a Conjugate Gradient algorithm with re-orthogonalization (Rödenbeck 2005).

4.2.2 Characteristics of the inversion set up

4.2.2.1 A-priori information and uncertainties

Table 4.1 Optimized VPRM parameters SW_0, λ_{SW}, α, β for different vegetation classes[a]

	SW_0	λ_{SW}	α	β
Evergreen forest	275	0.226	0.288	-1.10
Deciduous forest	254	0.215	0.181	0.84
Mixed forest	446	0.163	0.244	-0.49
Open shrub	70	0.293	0.055	-0.12
Crop	1132	0.086	0.092	0.29
Grass	528	0.119	0.125	0.017

[a]Units are as follows: SW_0 : W m^{-2}; λ_{SW}: µmole CO_2 $m^{-2}s^{-1}$ / (W m^{-2}); α: µmole CO_2 $m^{-2}s^{-1}$ / ^0C; β: (µmole CO_2 $m^{-2}s^{-1}$).

The a-priori CO_2 flux fields were derived from the Vegetation Photosynthesis and Respiration Model, VPRM (Mahadevan et al., 2008). VPRM uses ECMWF operational meteorological data for radiation (downward shortwave radiative flux) and temperatures (T2m), the SYNMAP landcover classification (Jung et al., 2006), and EVI (enhanced vegetation index) and LSWI (land surface water index) derived from MODIS (Moderate Resolution Imaging Spectroradiometer). Model parameters were re-optimized for Europe using eddy covariance measurements made during 2007 from 47 sites (a full site list is given in Kountouris et al. (2015); we excluded some sites due to insufficient temporal data coverage or lack of representativeness). To mediate the impact of data gaps, a data density weighting was introduced that takes into account the coverage of different times of the day (using 3-hour bins) in the different seasons. Optimized parameters are shown in Table 4.1. The net ecosystem exchange at hourly scale and at 0.25° x 0.25° spatial resolution for 2007 was simulated with the optimized parameters for the European domain shown in Figure 4.1. The domain-wide aggregated biospheric carbon budget for 2007 derived that way from VPRM was found to be -0.96 GtC y^{-1} (i.e. uptake by the biosphere). Note that without the density weighting an even stronger flux of -1.35 GtC y^{-1} was derived, indicating the importance of proper treatment of data gaps by either gap-filling or by the inclusion of weights.

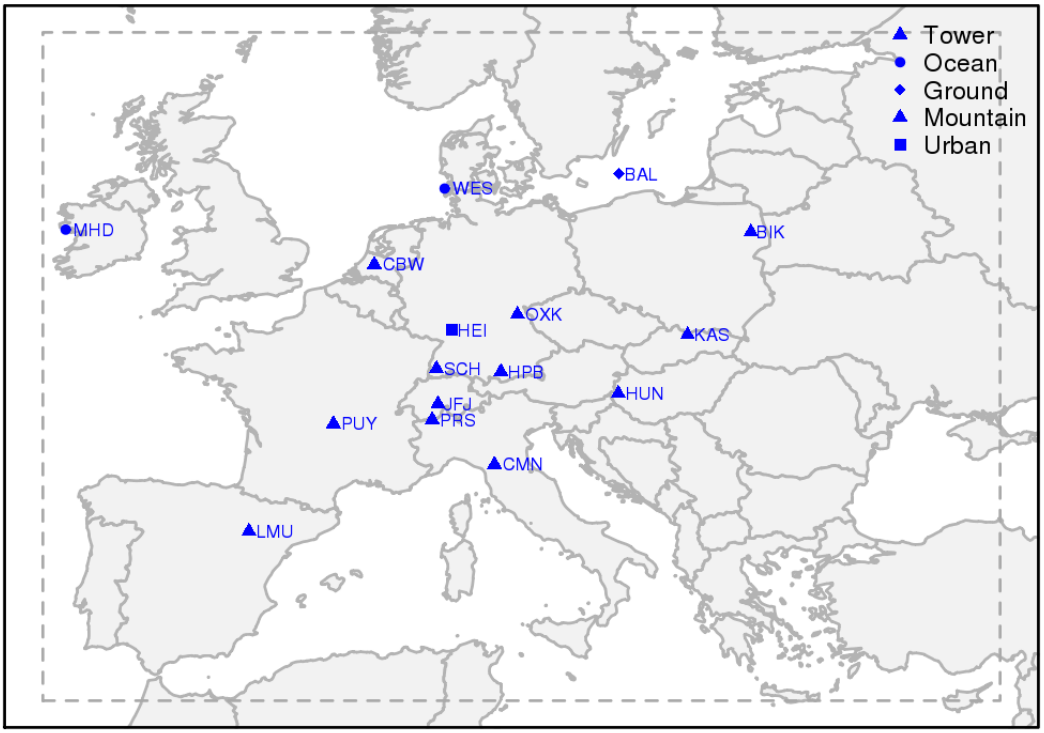

Figure 4.1 Domain of the inversions (dashed rectangle). Locations of the atmospheric measurement stations are shown with blue marks. Red stars denote the eddy covariance locations used for flux comparisons at grid scale.

Additionally, biogenic CO_2 fluxes were simulated with the BIOME-BGC model, specifically its global implementation as GBIOME-BGCv1 (Trusilova and Churkina 2008) at the same 0.25°x 0.25° spatial and hourly temporal resolution. The purpose of the second flux field is to provide a perfectly known flux distribution as "true" fluxes that can be used to generate synthetic observations. The a-priori flux in a real-data inversion would have three components including fossil fuel and ocean fluxes

$$f_{pr} = f_{pr,nee} + f_{pr,ff} + f_{pr,oc} \qquad \textbf{4.5}$$

We note that for the synthetic case the last two terms are set to zero. Similarly the deviation term (the data-derived correction to the a-priori fluxes) of the flux model consists of the terms referring to NEE, fossil fuel, and ocean fluxes but equivalently the last two terms are set to zero for the synthetic inversion.

$$F\delta s = (F_{nee}, F_{oc}, F_{ff}) \begin{pmatrix} \delta s_{nee} \\ \delta s_{oc} \\ \delta s_{ff} \end{pmatrix} \qquad \textbf{4.6}$$

Note that the a-priori error covariance matrix does not explicitly appear in the inversion, but is included though the second term in Eq. 4.8 (see section 4.2.2.2).

According to this formulation the columns of G_{tcor} and G_{xycor} contain the spatiotemporal extents of the individual NEE pulses (range of values between 0 and 1) and the diagonal matrix $f_{sh}(x,y,t)$ contains the pixel-wise a priori uncertainties. These uncertainties were chosen to be flat (constant) in space and time. For more detailed information the reader is referred to Rödenbeck et al. (2005).

The total prior uncertainty was chosen according to the mismatch between VPRM and BIOME-BGCv1, calculated as the annual and domain wide integrated flux mismatch. Prior fluxes and the fluxes representing the synthetic truth are strongly different (-0.96 GtC y^{-1} and -0.31 GtC y^{-1} for VPRM and GBIOME-BGCv1, respectively). The error structure used for the synthetic study is estimated according to the method applied in Kountouris et al. (2015). Time-series of daily fluxes were extracted for both biosphere models at grid cell locations where an EC station exists. Then spatial and temporal autocorrelation analysis was performed on the daily model-model flux residuals, yielding a spatial correlation length scale of 566 km and a temporal correlation scale of 30 days.

The eddy covariance station locations used for this analysis were exactly the same as in Kountouris et al. (2015) ensuring similarity in the derivation of the error structure for the synthetic data inversions. However of note is that for the synthetic data inversions, prior fluxes from VPRM model were not optimized against GBIOME-BGCv1 "true" fluxes.

The implicitly defined prior error covariance matrix contains diagonal elements of (1.45 μmol m^{-2} s^{-1})^2, which reflect the variance from model-model flux mismatches at the 50 km spatial resolution of the state space. Exponentially decaying spatial correlations were implemented with a correlation scale of 766 km at the zonal and 411 km at the meridional direction, roughly corresponding to the 566 km correlation scale yielded from the model-model residual autocorrelation analysis and preserving the same zonal/meridional ratio as in the global inversion. Temporal autocorrelation was set to 31 days, which is consistent with the Kountouris et al. (2015) analysis. These scales result in an uncertainty for the spatiotemporal component (E_{st}) domain-wide and annually integrated of 0.44 GtC y^{-1}. We chose two different approaches to increase the prior uncertainty at domain-wide and annually integrated scale such that it matches the mismatch of 0.65 GtC y^{-1} between the two biosphere models. First we inflate the error by scaling the error covariance matrix, this case is referred to as base case B1 hereafter. The second approach, referred to as scenario S1, could be considered as a more formal way: we introduce an additional degree of freedom to the inversion system by allowing for a bias term. This term is spatially distributed according to the annually averaged VPRM respiration component, and is kept constant in time. The error E_{BT} of the bias component was adjusted such that the total prior error E_{tot} for annually and domain-wide integrated fluxes matches the targeted total uncertainty:

$$E_{tot}^2 = E_{ST}^2 + E_{BT}^2 \qquad \textbf{4.7}$$

This resulted in an overall uncertainty E_{tot} of 0.65 GtC y^{-1}, which is identical to the mismatch between the two biosphere models.

4.2.2.2 State space

The inversion system optimizes additive corrections to three-hourly fluxes in a sense that the posterior flux estimate can be given by the sum of a fixed a priori term (first term of the right hand side in Eq. 4.8) and an adjustable term (second term in Eq. 4.8). The latter has a-priori a zero mean and unit variance. The biogenic fluxes can be defined as follows:

$$f(x,y,t) = f_{fix}(x,y,t) + f_{sh}(x,y,t) \cdot \sum_{m_t}^{N_t} \sum_{m_s}^{N_s} G_{tcor,m_t}(t) \cdot G_{xycor,m_s}(x,y) \cdot p_{inv,m_t,m_s} \quad \textbf{4.8}$$

where f_{sh} is a shape function which defines the adjustable term. The spatial and temporal correlation structures of the uncertainty are described by the pulse response functions G_{xycor} and G_{tcor} respectively. The term p_{inv} contains the adjustable parameters which they a-priori have, a Gaussian distribution with zero mean and unit variance.

For the S1 case the posterior flux estimates can be derived by adding the optimized bias flux field to Eq. 4.8

$$f(x,y,t) = f_{fix}(x,y,t) + f_{sh}(x,y,t) \cdot \sum_{m_t}^{N_t} \sum_{m_s}^{N_s} G_{tcor,m_t}(t) \cdot G_{xycor,m_s}(x,y) \cdot p_{inv,m_t,m_s} + \\ + f_{sh}^{BT}(x,y) \cdot \sum_{m_t}^{N_t} G_{tcor,m_t}(t) \cdot p_{BT} \quad \textbf{4.9}$$

The bias term f^{BT} follows a flux shape (here we used annually averaged respiration, with no temporal variation).

4.2.2.3 Observation vector and uncertainties

Table 4.2 Information on the stations used for the regional inversions. Same network applied for the synthetic, and the real data inversions in Kountouris et al. (2016). In first column the term "type" stands for continuous (C) or flask (F) data.

Site Code / type	Name	Latitude (^{o})	Longitude (^{o})	Height (m.a.s.l.) (m)	Measurement height (above ground) (m)	Model height
BAL/F	Baltic Sea, Poland	55.50	16.67	8	57	28
BIK/C	Bialystok, Poland	53.23	23.03	183	90	90

CBW/C	Cabauw, Netherlands	51.58	4.55	-2	200	200
CMN/C	Monte Cimone, Italy	44.18	10.7	2165	12	670
HEI/C	Heidelberg, Germany	49.42	8.67	116	30	30
HPB/F	Hohenpeissen-berg, Germany	47.80	11.01	934	50	10
HUN/C	Hegyhatsal, Hungary	46.95	16.65	248	115	96
JFJ/C	Jungfraujoch, Switzerland	46.55	7.98	3572	10	720
KAS/C	Kasprowy Wierch	49.23	19.93	1987	5	480
LMU/C	La Muela, Spain	41.36	-1.6	570	79	80
MHD/C	Mace Head, Ireland	53.33	-9.90	25	10	15
OXK/C	Ochsenkopf, Germany	50.03	11.81	1022	163	163
PRS/C	Plateau Rosa, Italy	45.93	7.71	3480	-	500
PUY/C	Puy De Dome, France	45.77	2.97	1465	10	400
SCH/C	Schauinsland, Germany	47.92	7.92	1205	-	230
WES/C	Westerland, Germany	54.93	8.32	12	-	15

The observation vector c_{meas} contains mixing ratio observations at all site locations and sampling times. A common procedure to derive synthetic observations is to create a “true” flux field by adding some error realizations to the a-priori fluxes (Schuh et al., 2009; Broquet et al., 2011) or

to perturb the resulting synthetic observations (Wu et al., 2011). For the current study instead we use a different biosphere model, the GBIOME-BGCv1 model, to derive biogenic CO_2 fluxes at hourly scale. Then a forward transport model run was performed to create synthetic mixing ratios at hourly resolution for each station location. This choice of using two different biosphere models for deriving the a-priori and the "true" fluxes is expected to increase the realism of the synthetic data study, given the fact that the real spatiotemporal flux distribution is highly unknown (though the model-to-model difference may not accurately reflect the model errors either). For the synthetic study, observations were created for the same station locations and observation times as in the real observation time series which are used in the second part of this study (Kountouris et al., (2016)). An overview of the atmospheric stations is given in Table 4.2. The data coverage per station is shown in Figure 4.2. Only daytime observations were considered (11:00 – 16:00 local time) since the transport model is expected to perform worse during night when a stable boundary layer forms. An exception is made for mountain stations that measure the free troposphere, where only nighttime observations (23:00 – 04:00 local time) were considered, as this time can be better represented by the transport model. In total 20273 hourly observations from the year 2007 were used.

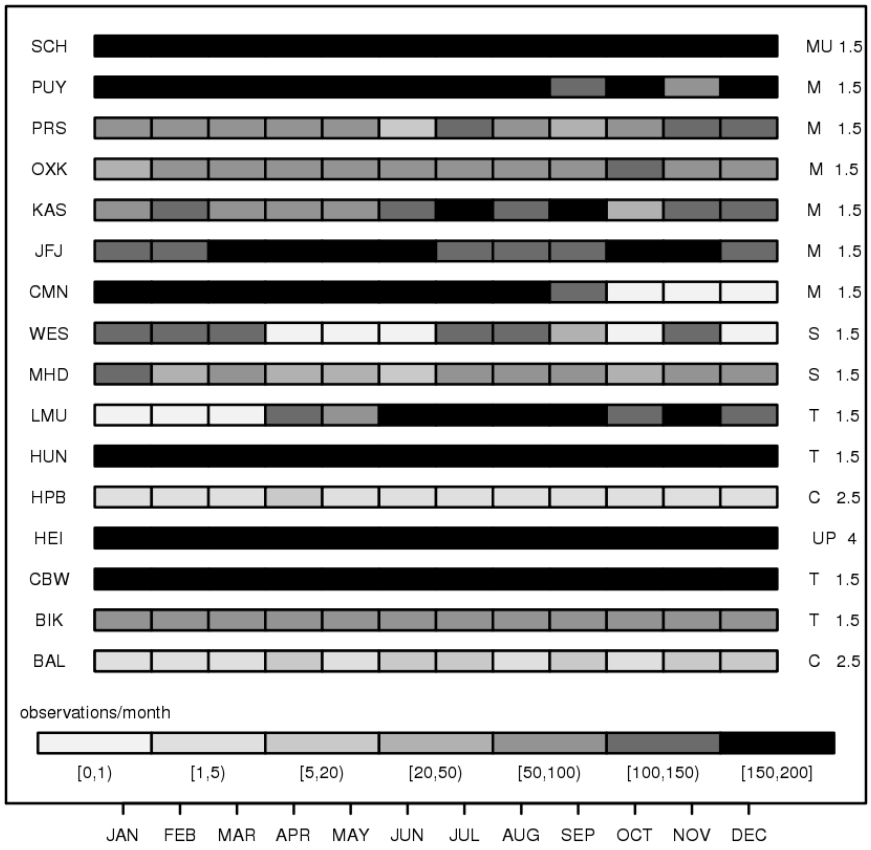

Figure 4.2 Monthly data coverage plot for the atmospheric stations used in the regional inversions. Left column shows the code name and the right columns show the station class and the assigned uncertainty in units of ppm. "C" stands for continental sites near the surface, "T" for continental tall towers, "S" for stations near shore, "M" for mountain sites, "MU" for mountain sites with diurnal upslope winds and "UP" for urban pollutant.

The model-data mismatch uncertainty associated with each measurement is expressed as a diagonal covariance matrix, and contains measurement errors and errors from different components describing the modeling framework (i.e. model errors due to imperfect transport,

aggregation errors, etc.) (Gerbig et al., 2003b). For the current study, all sites are classified according to their characteristics (e.g. tall tower, mountain sites etc.), and uncertainties were defined depending on the site class (Figure 4.2, legend on the right). The uncertainties are considered as representative for current inverse modeling systems. Although the measurement error covariance is a diagonal matrix, we do consider for temporal correlations via a data density weighting (see Section 4.2.1).

4.2.2.4 Atmospheric transport

For the synthetic data study only the regional atmospheric model STILT was used to create the observations with a forward run, and to perform the inversion. This was feasible since the synthetic CO_2 observations are only influenced by fluxes occurring within the DoI, hence global runs to retrieve boundary conditions at the edge of DoI are not necessary. The transport matrix for the regional inversions was generated in form of pre-calculated footprints (sensitivities of atmospheric observations to upstream fluxes) at 0.25 degrees spatial and hourly temporal resolution for the full year 2007. STILT trajectory ensembles were driven by ECMWF meteorological fields (Trusilova et al., 2010), and computed for 10 days backwards in time, ensuring that nearly all trajectories have left the domain of interest.

4.2.3 Metrics for performance evaluation

Following Rödenbeck et al. (2003) we evaluate the goodness of fit for each station (station specific χ^2). The modeled dry mole fractions should be with 68% probability within the ± 1σ range from the observed mole fractions. This is equivalent to the requirement that the dry mole fraction part of the cost function defined as the sum of hourly squared differences, divided by the uncertainty interval and the number of observations n (Eq. 4.10), should be close to unity.

$$\chi_c^2 = \frac{\sum_t \frac{(\Delta c_t)^2}{\sigma_t^2}}{n} \qquad \textbf{4.10}$$

Another important aspect is the reduced χ_r^2 metric that compares the a-priori model performance with the specified error structure by dividing the squared residuals of optimized minus observed dry mole fractions by the squared specified uncertainties. This is also equivalent to two times the cost function at its minimum divided by the number of degrees of freedom (effective number of observations) (Thompson et al., 2011):

$$\chi_r^2 = 2\frac{J_{\min}}{n} \qquad \textbf{4.11}$$

Again, a correct balance should be close to unity. Smaller values suggest that the model performance was better than specified in the covariance structure and hence the assumed uncertainties (denominator) were conservative.

In flux space, we evaluate the inversion performance, by comparing the retrieved flux estimates against the synthetic fluxes ("true") at different temporal and spatial scales: annually and monthly integrated fluxes, domain-wide and at country scale. In particular we are interested in capturing the "true" fluxes down to country scale. For that we assess monthly posterior retrievals which we compare to reference data ("true" fluxes), country aggregated, using a Taylor diagram. This diagram provides a concise statistical summary of how well patterns match each other in terms of their correlation and the ratio of their variances.

4.3 Results

The purpose of the synthetic study is to evaluate the system set-up with a realistic approach. To evaluate the ability of the system to retrieve the synthetic true fluxes we visualize spatially distributed fluxes and we study spatially integrated (domain and national scale) as well as temporally (annual and monthly scale) integrated fluxes.

4.3.1 CO_2 mole fractions

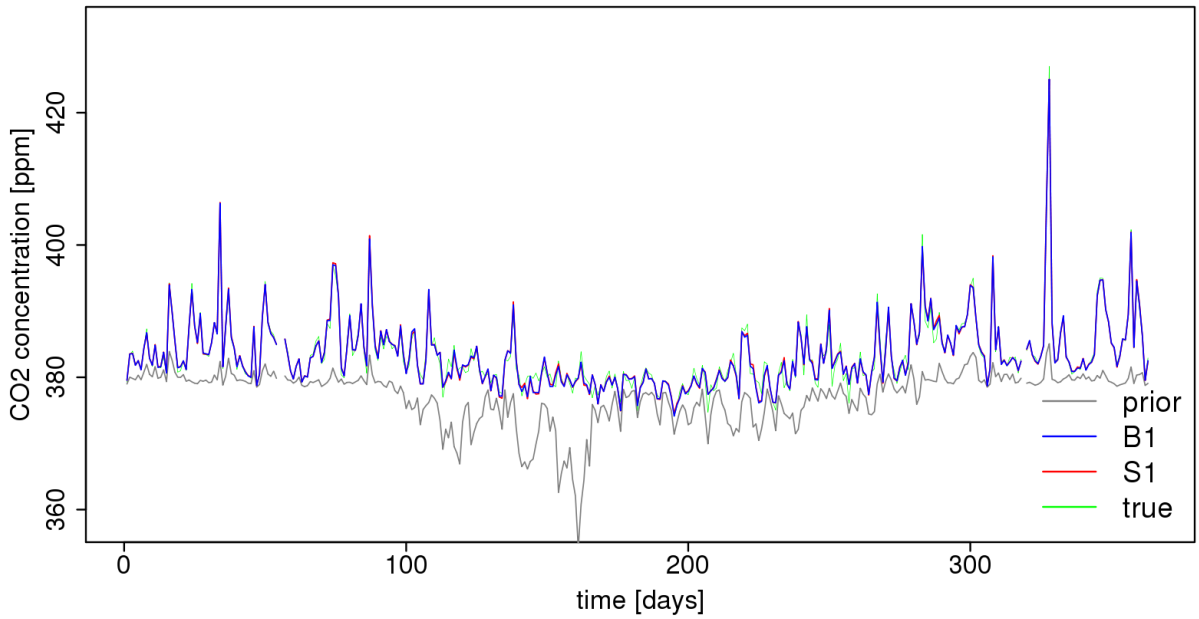

Figure 4.3 Daily nighttime (23:00-4:00 UTC) averages for prior, true, and posterior CO_2 dry mole fraction time series for the mountain site Schauinsland. Time starts at 1st January 2007.

A comparison of true and modeled CO_2 dry mole fractions from forward runs of the optimized fluxes can reveal the goodness of fit, realized through the optimization process. Such a comparison is presented in Figure 4.3 for the Schauinsland (SCH) continuous station. Both B1 and S1 inversions significantly reduce the misfit between the synthetic (truth) and the a-priori mole fractions. The RMSD between the prior/posterior from the "true" timeseries for all stations (Table 4.3) shows an average reduction of around 74% and 76% for the S1 and B1 inversions respectively.

Table 4.3 RMSD (first column in ppm) and correlation coefficients (second column) between known truth and prior/posterior CO_2 dry mole fractions for daily "daytime" or "nighttime" averaged values and for each station. The third column shows χ^2, the normalized dry mole fraction mismatch per degree of freedom for 7-day averaged residuals, as a measure of how well the data were fitted. The format for each station is as follows: RMSD | r^2 | χ^2.

	Prior	B1	S1
BAL	4.78 | 0.07 | 18.44	0.89 | 0.97 | 0.48	1.02 | 0.96 | 0.37
BIK	5.28 | 0.43 | 15.50	1.20 | 0.97 | 0.18	1.29 | 0.97 | 0.25
CBW	8.60 | 0.04 | 74.29	0.99 | 0.99 | 1.31	1.06 | 0.99 | 1.34
CMN	2.68 | 0.33 | 6.31	0.74 | 0.93 | 0.08	0.78 | 0.92 | 0.10
HEI	11.39 | 0.37 | 12.97	1.83 | 0.98 | 0.36	1.84 | 0.98 | 0.37
HPB	7.73 | 0.35 | 26.58	1.01 | 0.99 | 0.21	1.19 | 0.99 | 0.31
HUN	6.50 | 0.63 | 31.89	1.36 | 0.98 | 0.21	1.46 | 0.98 | 0.25
JFJ	3.12 | 0.21 | 3.93	1.24 | 0.86 | 0.24	1.31 | 0.84 | 0.27
KAS	4.00 | 0.32 | 10.67	0.73 | 0.98 | 0.11	0.80 | 0.97 | 0.15
LMU	3.42 | 0.19 | 6.5	0.79 | 0.95 | 0.12	0.86 | 0.94 | 0.16
MHD	1.53 | 0.0002 | 0.83	0.65 | 0.09 | 0.16	0.68 | 0.06 | 0.17
OXK	6.10 | 0.21 | 38.50	3.35 | 0.76 | 0.76	3.40 | 0.75 | 0.80
PTR	2.32 | 0.15 | 2.46	0.70 | 0.92 | 0.30	0.74 | 0.91 | 0.33
PUY	4.27 | 0.15 | 12.06	0.68 | 0.97 | 0.06	0.73 | 0.15 | 0.09
SCH	4.76 | 0.26 | 21.17	0.90 | 0.97 | 0.07	0.95 | 0.97 | 0.09

Prior correlations (prior vs. true dry mole fractions), have an averaged value of 0.46 which is increased to 0.93 for both inversions. Significant differences between the two inversions were not found apart from a slightly larger decrease of the RMSD for the B1 case. Figure 4.1 summarizes the capability of the inversions to capture the true signal at each station location in form of a Taylor diagram, indicating that the inversions showed a significant increase of the correlation for all sites. Further the variance of the modeled time-series is significantly closer to the variance of the true signal.

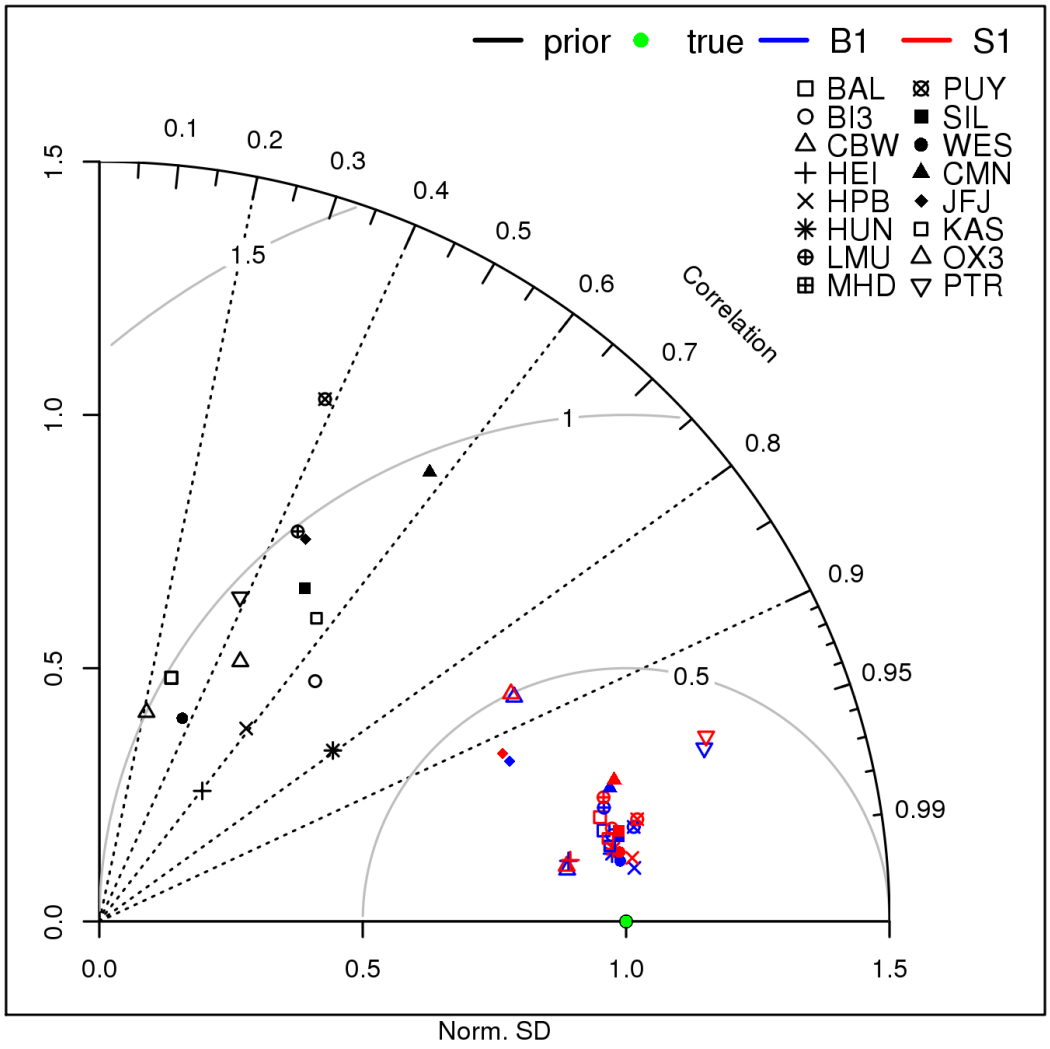

Figure 4.4 Taylor diagram for modeled and measured time-series of CO_2 dry mole fractions. Prior (black), true (green, the perfect match of modeled and true time-series) and the different inversion cases (R0 blue; R1 red) are displayed. Different symbols denote different atmospheric stations. The normalized SD was calculated as the ration of the SD of the modeled time-series to the SD of observations.

To estimate the goodness of fit we consider the station specific χ_c^2 values (Eq. 4.10), using here 7-day aggregated residuals instead of hourly to match the temporal scale of one week of the observation error. Values smaller than 1 are found for most of the stations with a mean value of 0.28 and 0.32 for the B1 and S1 cases respectively, suggesting a good fitting performance for all stations and for both inversions. The results are comparable with those found in the Rödenbeck et al. (2003) study. The reduced chi-squared (Eq. 4.11) was found to be 0.21 for both cases, indicating that the error variance is overestimated making the error assumption rather conservative.

4.3.2 Flux estimates and uncertainties

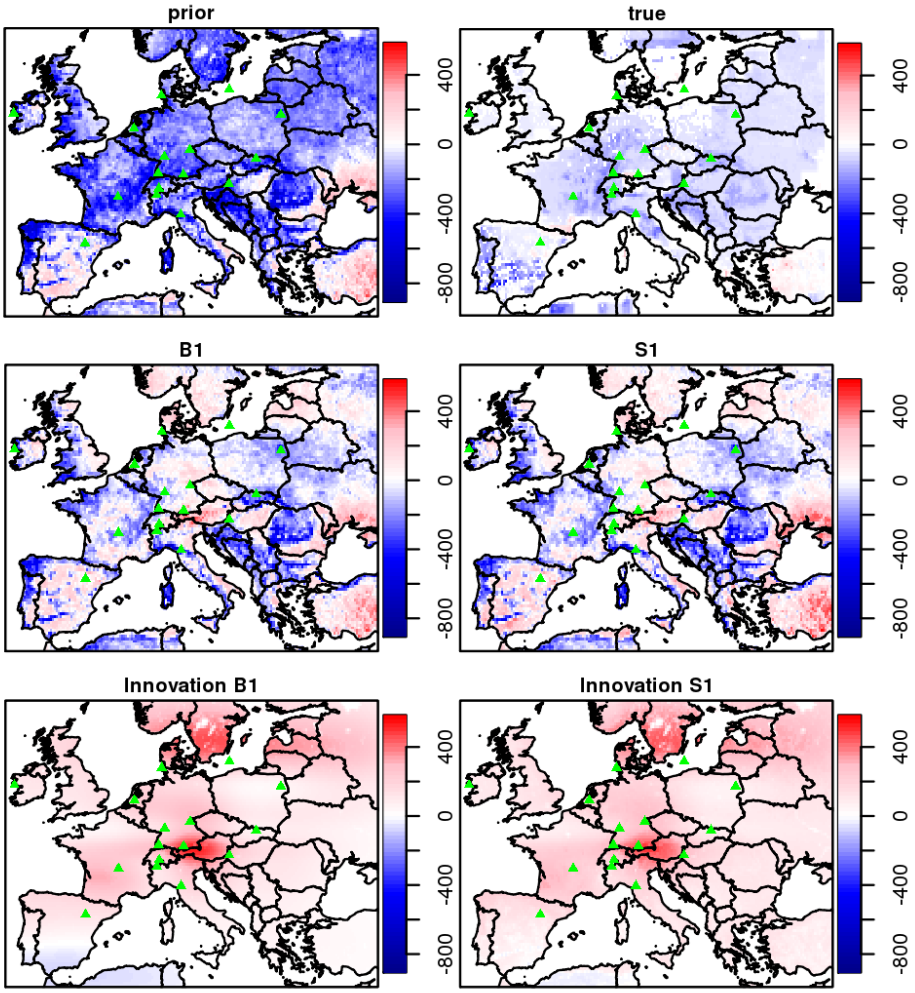

Figure 4.5 Annual spatial distribution for the prior, true, and posterior biogenic flux estimates for the two synthetic inversions S1 and B1 (top two rows), and flux innovation defined as the difference posterior - prior (bottom row). Fluxes are given in units of $gCy^{-1}m^{-2}$.

The spatial distributions of the annual biosphere-atmosphere exchange fluxes for the prior, the known truth, and the posterior cases are presented in Figure 4.5. Note that annual fluxes between the two biosphere models used for prior fluxes and true fluxes are substantially different. The inversion significantly adjusts the spatial flux distribution mainly in central Europe, where a denser atmospheric network exists. The absolute annual mean difference in fluxes (|mean(true – prior)| and |mean(true – posterior)|) is greatly reduced from 70.8 $gCm^{-2}y^{-1}$ to 14.7 $gCm^{-2}y^{-1}$ and 24.6 $gCm^{-2}y^{-1}$ for the B1 and S1 inversions respectively. Detailed patterns, however, are not well reproduced: the fraction of explained spatial variance in the true fluxes (measures as squared Pearson correlation coefficient) decreases from the prior (0.17) to the posterior (0.07 and 0.06 for the cases B1 and S1, respectively). When evaluating this at monthly scales, the fraction of explained spatial variance increases in the posterior estimates compared to the prior for winter months from around 0-15% to about 15-50%, while during the growing season typically a decrease from around 10-35% to about 0-34% is found.

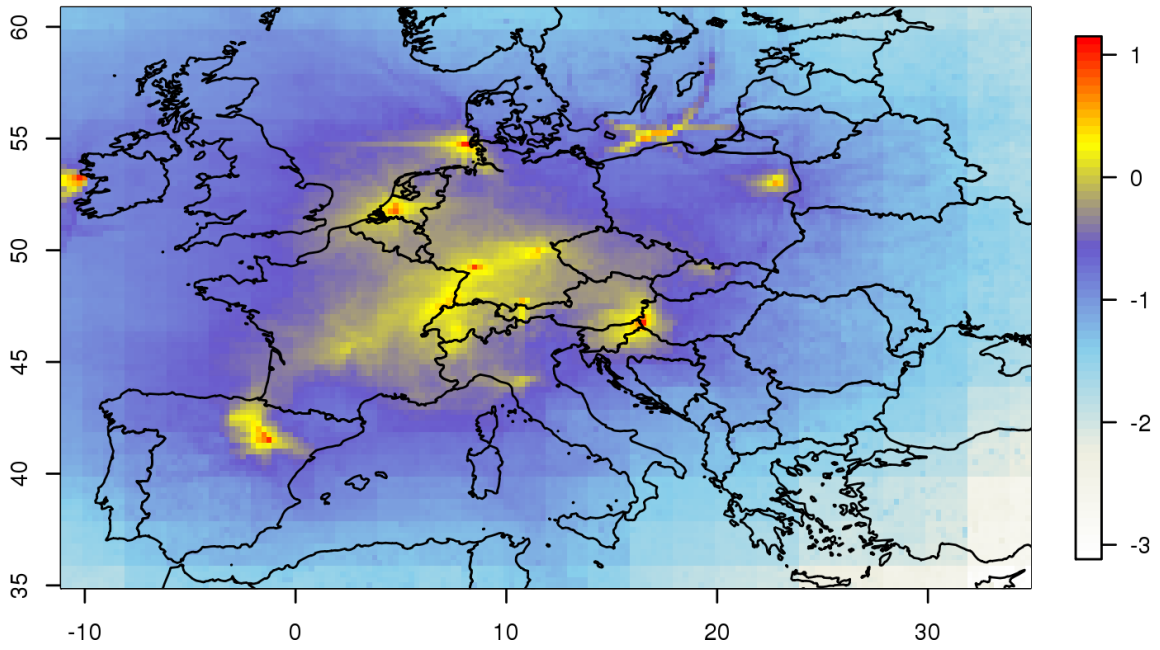

Figure 4.6 Annual integrated influence for 2007 of the current atmospheric network. Footprint influence is presented in a logarithmic scale and units are in $\log_{10}$[ppm/(μmol/m^2/s)]

The accumulated footprint of the atmospheric network is shown in Figure 4.6, clearly indicating the strongest constraint on fluxes in central Europe. Interestingly both error structures from S1 and B1 inversions produce posterior fluxes that have approximately the same spatial distribution. When separating the spatiotemporal component from the bias component (in S1 case) we can identify differences between the two inversions. Significant deviations of the spatial flux distribution between the spatiotemporal components were found: The spatiotemporal component in the S1 case has a domain wide annual flux correction of 0.39 GtC y^{-1} (prior – posterior) while the corresponding term in the B1 case has a correction of 0.78 GtC y^{-1}. Nevertheless standard deviations of the corrections with respect to the true spatial flux distribution (true – posterior) found to have no significant difference ($6.88*10^{-5}$ and $7.38*10^{-5}$ GtC y^{-1}cell^{-1} for S1 and B1 respectively). We do not observe any strong correction in the south-eastern part of Europe as it cannot be "seen" from the atmospheric network due to the distance to the observing sites and the prevailing westerly winds. This could also be inferred from the flux innovation plots (see Figure 4.5) defined as the difference between prior and posterior fluxes. Only very small or even no corrections occurred in this area.

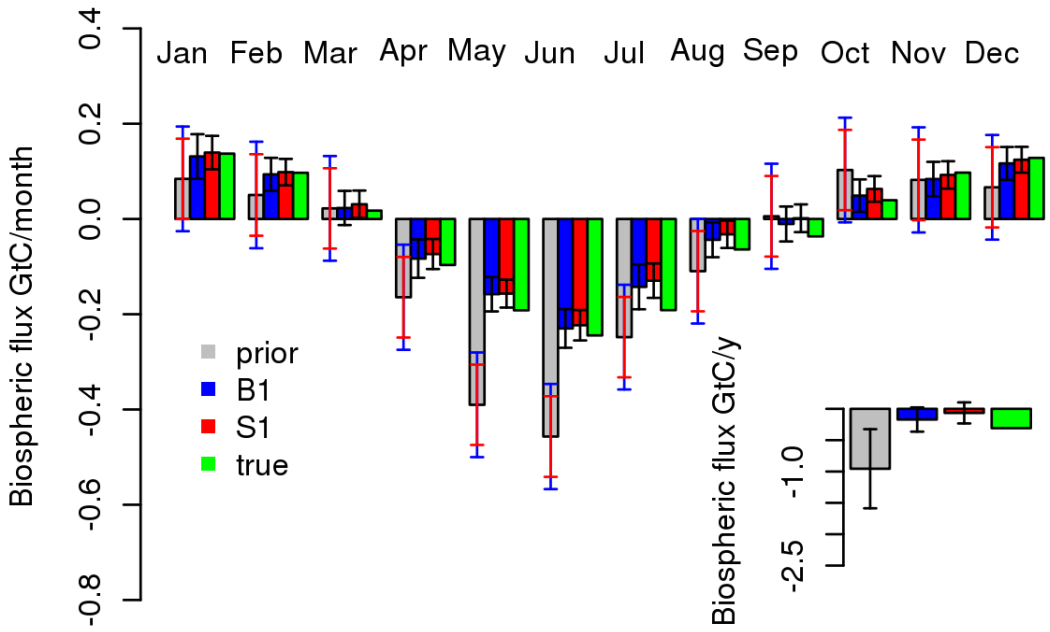

Figure 4.7 Monthly and annual carbon flux budget, integrated over the European domain. Note that both inversions share the same annual prior uncertainty but monthly uncertainties differ. Blue and red error bars denote the prior uncertainty for the B1 and S1 scenarios respectively.

We are specifically interested in the ability of the inversion system to capture integrated fluxes over time and space. Figure 4.7 shows an overview of the domain-integrated fluxes at a monthly and annual scale. Despite the remarkably larger a-priori (VPRM) sink compared to the synthetic truth (GBIOME-BGCv1) during the growing season, both inversions, with and without the bias term, produce posterior flux estimates that fully capture the "true" monthly and annually integrated fluxes. While the monthly posterior estimates give no clear evidence on which inversion performs better, retrievals at annual scale slightly favor the inversion without the bias term (B1 case). A difference was observed in the prior uncertainties between the two inversions. While both were scaled to have the same prior annual uncertainty, the B1 inversion has systematically larger prior monthly uncertainties than the S1 as a result of the inflated spatiotemporal component of the prior error covariance. Posterior uncertainties were found to be similar, and include or are close to including (S1 case) the true flux estimates. The uncertainty reduction for annually and domain-wide integrated fluxes, defined as the difference between prior and posterior uncertainties normalized by the prior uncertainty, was found to be 73% and 69% for the S1 and B1 respectively.

Table 4.4 Performance of the two error structures expressed as the spatial RMSD of the optimized monthly and annual NEE fluxes compared to the truth for the whole domain in μmole m^{-2} s^{-1}.

	Annual	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
prior	0.38	0.61	0.53	0.55	1.06	1.26	1.56	1.17	0.94	0.65	0.57	0.63	0.63
B1	0.33	0.46	0.40	0.45	0.84	0.99	1.21	1.00	0.86	0.63	0.43	0.46	0.44
S1	0.34	0.48	0.41	0.45	0.86	1.01	1.24	1.03	0.86	0.63	0.45	0.47	0.45

In order to assess how well the posterior estimates agree with the true fluxes, root mean square difference (RMSD) between true and posterior monthly integrated gridded fluxes were computed (Table 4). Both inversions B1 and S1 show a similar reduction in the RMSD values compared to the prior. The same picture emerges for the annually integrated fluxes.

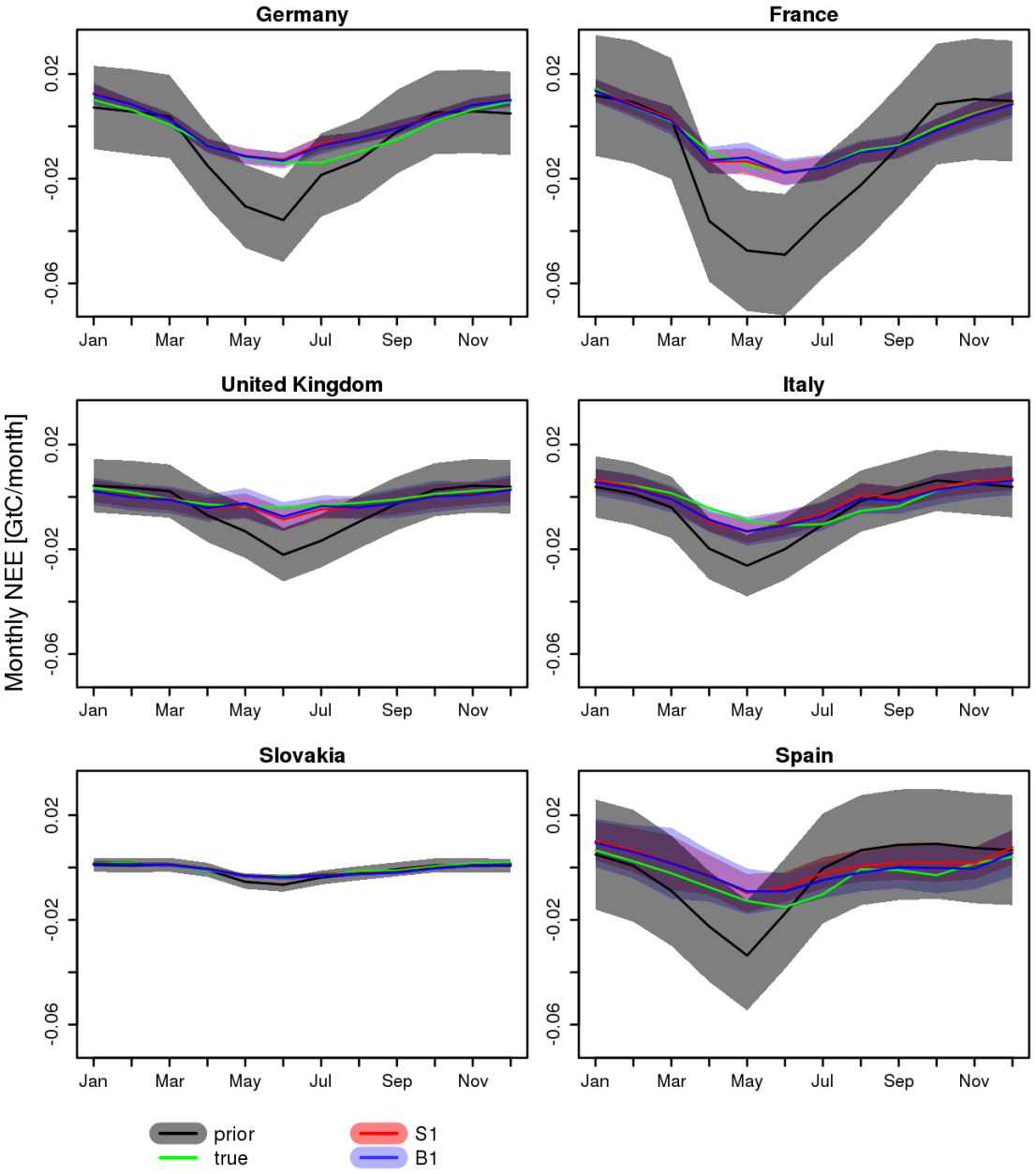

Figure 4.8 Temporal evolution of monthly NEE for selected European countries for the synthetic data inversion.

Of particular interest is the performance of the system at regional scale, specifically at national level. Figure 4.8 shows monthly fluxes for selected European countries, including the prior, true and posterior estimates with the corresponding uncertainties. Both error structures show a similar performance. Despite the large prior misfit, the system succeeded in retrieving monthly fluxes at country level. Better constrained regions mainly located in central Europe show the ability to broadly capture the temporal flux variation at monthly scale. Figure 4.9 summarizes in a Taylor diagram the inversion performance for each EU-27 country, showing the improvement of monthly and country aggregated fluxes (perfect match would be if the head of the arrow coincides with the reference point marked as green bullet). It is worth mentioning that also for

regions that are less constrained by the network, such as Great Britain, Spain, Poland and Romania, the inversions still improved the posterior estimates compared to the prior estimates (see also Figure 4.9).

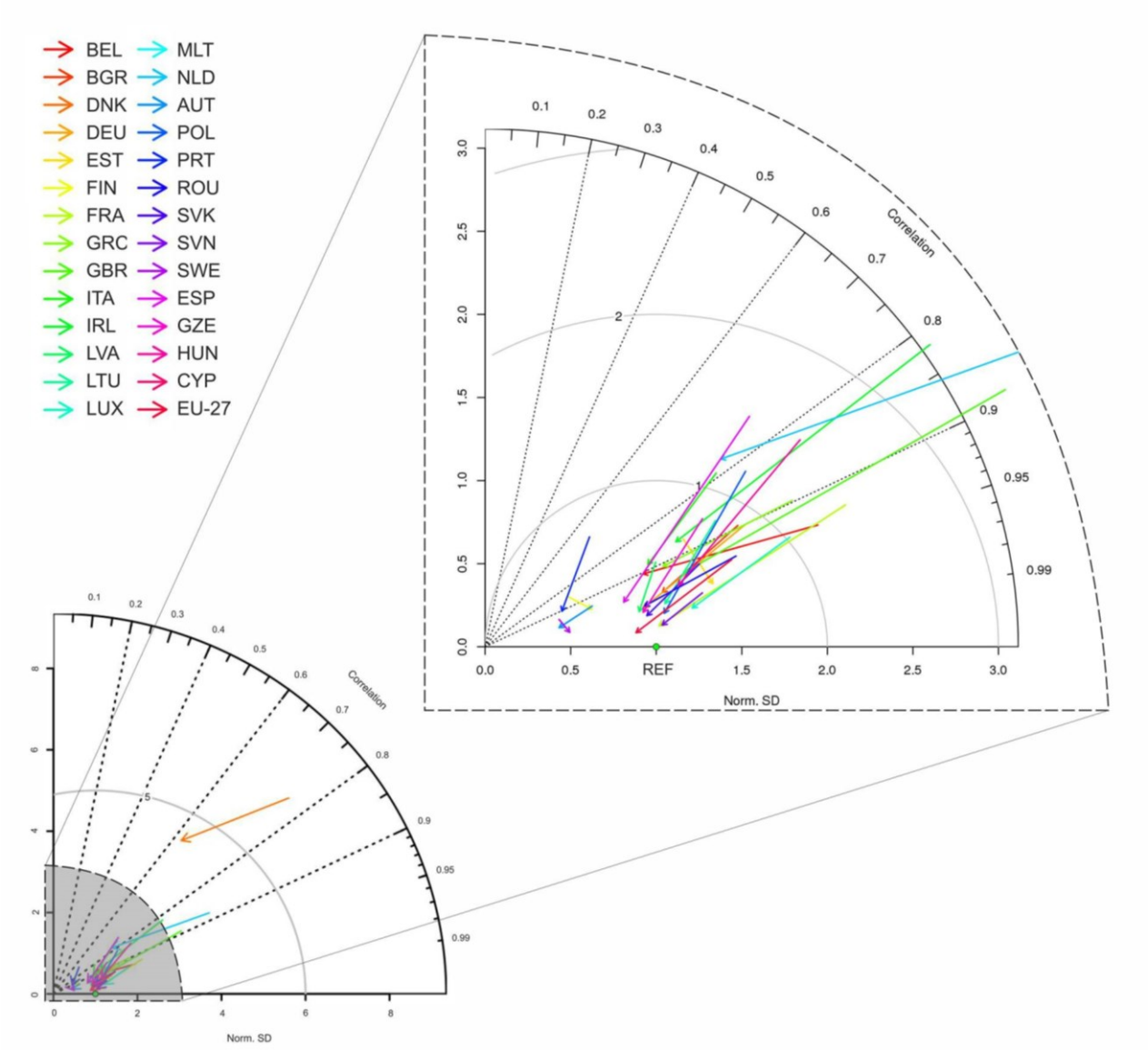

Figure 4.9 Overview of the model performance summarized in a Taylor diagram. Posterior and prior monthly and country scale aggregated biospheric fluxes are compared against the reference fluxes ("true"). Each line corresponds to a different country. The starting point of each arrow shows prior/reference comparison and the ending point the posterior/reference comparison. Ideally the ending point should coincide with the green point which represents the reference model.

4.3.3 Evaluation with synthetic eddy covariance data

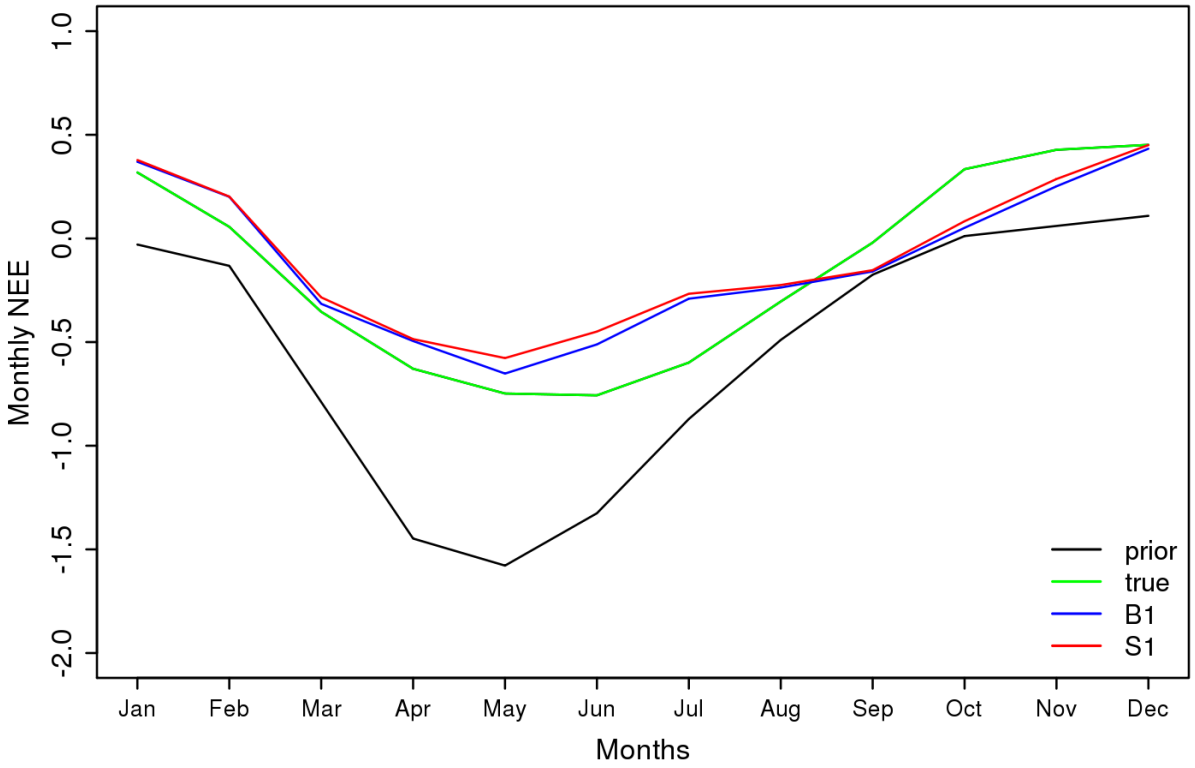

Figure 4.10 Mean monthly NEE averaged over the 53 different eddy covariance site locations as reported in Kountouris et al. (2015). A priori (black), true (green), and posterior fluxes for scenarios B1 (blue) and S1 (red) are shown. Units are in $gCm^{-2}day^{-1}$.

In order to investigate the potential of using eddy covariance measurements for evaluating the retrieved CO_2 fluxes, monthly fluxes from the prior (VPRM), the truth (GBIOME-BGCv1), and the posterior for cases B1 and S1 were extracted at the grid cell locations where eddy covariance stations exist, using the same 53 sites as in Kountouris et al. (2015). The corresponding fluxes were then aggregated over all sites, using a weight that compensates for the asymmetry between number of flux towers for specific vegetation types and the fraction of land area covered by the specific vegetation type. Prior fluxes show a systematically larger uptake compared to the truth, predominantly during the growing season with maximum differences of 0.8 gCm^2day^{-1} (Figure 4.10). Posterior estimates for both cases captured the magnitude of the true fluxes, with maximum differences of around 0.3 gCm^2day^{-1} during June/July. A significantly larger correction is apparent during spring and summer compared to winter and fall. The very close correspondence of these results with those shown in Figure 4.7 for the domain-wide monthly flux budget clearly shows that eddy covariance measurements can principally be used for validation of the inverse estimates at monthly timescales.

4.4 Discussion

4.4.1 Performance in flux space

Results from the synthetic experiment showed the strengths, but also the weaknesses of the system to retrieve the "true" spatial flux distribution. Although the error structure applied to this experiment was statistically coherent with the mismatch between prior and true fluxes, we note a limited ability of the current atmospheric network to retrieve fluxes at local scales. For coarser spatial scales (country level) the carbon budget estimates in the synthetic inversion showed a quite good performance at monthly and annual temporal scales. Further we observed an average reduction of the monthly uncertainties of 65% for the B1 case, and 64% for the S1 case. In combination with the fact that the flux estimates reproduce the "truth" within the posterior uncertainties, this gives us confidence in the accuracy of our estimates.

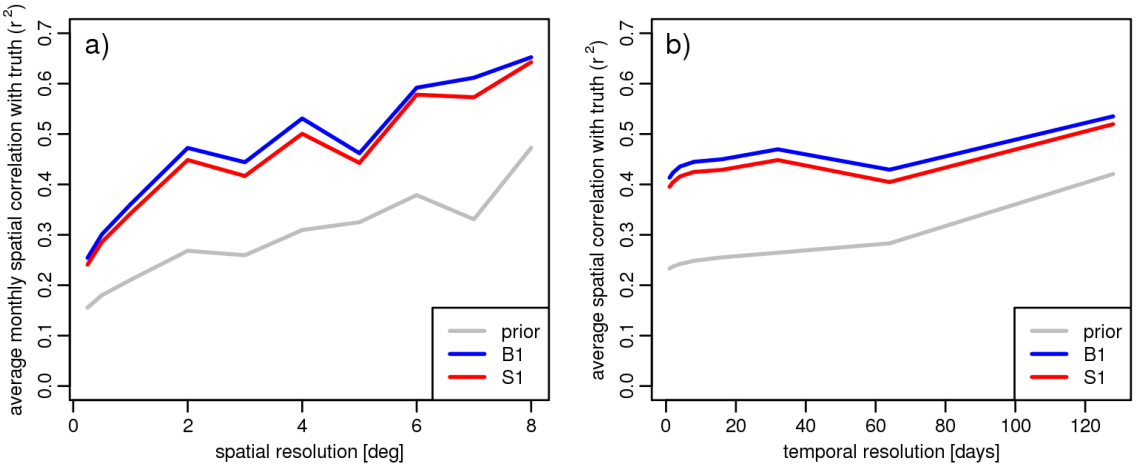

Figure 4.11 a): Mean spatial correlation of monthly fluxes with true fluxes as function of spatial flux aggregation scale for prior fluxes (grey), and for posterior fluxes from scenarios B1 (blue) and S1 (red). b): Mean spatial correlation of fluxes with true fluxes at 2 deg. spatial resolution as function of temporal flux aggregation scale for prior fluxes (grey), and for posterior fluxes from scenarios B1 (blue) and S1 (red).

Prior error correlation in time and space limits the scale, at which information can be retrieved from the inversion. The spatial correlation of several hundred kilometers implies that fluxes at scales smaller than this cannot be significantly improved by the inversion, as the results clearly showed. To assess this more quantitatively, the spatial correlation between a priori or retrieved and true monthly fluxes is calculated for different spatial aggregation scales (starting at 0.25 degree, fluxes were aggregated to 0.5, and then in 1-degree steps up to 8 degree). Results shown in Figure 4.11 a) indicate a nearly continuous increase of the spatial correlation of prior and posterior fluxes with increasing aggregation scale. The additional explained variance brought about by the inversion, i.e. the difference between posterior (red/blue line) and prior (grey line) flux correlation (r-square) with the truth, starts at low values around 0.1, and reaches values around 0.2 for scales larger or equal 2 degrees. Similarly, the spatial correlation between a priori and true fluxes for a given spatial aggregation of 2 degrees, but for different temporal aggregation scales ranging from 1 day to 128 days (Figure 4.11b) shows a continuous increase from about 0.23 to 0.42 (r-square), while the spatial correlation between retrieved and true fluxes only varies slightly between 0.4 and 0.53 (Figure 4.11b), red and blue lines). Here, the additional spatial variance explained by the retrieved fluxes is largest at around monthly time scales

(differences between prior and posterior r-square around 0.2), while at seasonal scales this additional explained variance is only around 0.1. Overall, this analysis confirms that there are preferred spatial and temporal scales at which the inversion retrieves the flux distribution best and where thus most information is gained. This is not dependent on whether or not a bias term is included in the state vector, as results for case B1 and S1 do not differ in this regard. It is important to realize that all other scales, at which the inversion does not provide much information, need to be properly represented by the a priori flux distribution. Thus the a priori fluxes need to be realistic at short spatial scales below about 200 km, at seasonal temporal scales, and of course at hourly time scales which are not retrieved by the inversion.

The annual spatial flux distribution of the B1 and S1 cases was found to be quite similar, indicating that inflating the uncertainty by a factor of 1.5 (B1 case, see also 4.2.2.1 section) or adding a bias component to compensate the inflation (S1 case) lead to a similar flux constraint. This could be explained due to the long correlation length (566 km) which drastically reduces the effective number of degrees of freedom, forcing the fluxes to be smoothly corrected, regardless the use of the bias component.

4.4.2 Performance in observation space

The high RMSD reduction in combination with the high correlation values and the captured variability between posterior and true dry mole fractions in the synthetic experiment suggest a good performance of the inversion system to retrieve the "true" mixing ratios. Nevertheless this is not surprising, as the atmospheric data are "fitted" by the inversion, and furthermore the forward and the inverse runs used identical transport, without any impact from imperfections in transport simulations.

The uncertainties in the flux space are statistically consistent with the model-model flux mismatch. However the reduced χ_r^2 values obtained from the inversions were rather small (around 0.21). This indicates that overall conservative uncertainties were assumed, and the small χ_r^2 values are a result from the assumed uncertainties in the observation space. Indeed uncertainties in the observation space include also transport uncertainties; however, given that the same transport is used to create synthetic observations and to perform the inversion, there is no actual model-data mismatch related to transport uncertainties, and so the assumed uncertainties are overestimated.

4.5 Conclusions

This paper describes the setup and the implementation of prior uncertainties as derived from model-eddy covariance data comparisons into an atmospheric CO_2 inversion. The inversion system assimilates hourly dry air mole fractions from 16 ground stations to optimize 3-hourly NEE fluxes for the study year 2007. Two different error structures were introduced to describe the prior uncertainty by either inflating the error or by adding an additional degree of freedom allowing for a long term bias. The need of this error inflation comes from the fact that the spatiotemporal model - data error structure alone underestimates prior uncertainties typically assumed for inversion systems at continental/annual scale. In this study we evaluate the Jena inversion system by performing a synthetic experiment and expanding the evaluation also to the

retrieved fluxes, whilst only the observation space was evaluated in Trusilova et al. (2010). Further we assess the impact when adding a bias term in the flux error structure. This study is a preparatory step to retrieve European biogenic fluxes using a data driven error structure consistent with model-flux data mismatches, which is described in the companion paper (Kountouris et al. 2016).

Significant flux corrections and error reductions were found for larger aggregated regions (i.e. domain-wide and countries), giving us confidence on the reliability of the results for a real data inversion. We found a similar performance for both error structures. A more detailed analysis of the spatial and temporal scales, at which the inversion provides a significant gain in information on the distribution of fluxes, clearly confirms that a) fluxes at spatial scales much smaller than the spatial correlation length used for the a prior uncertainty cannot be retrieved; b) the inversion performs best at temporal scales around monthly, and c) especially the small spatial scales need to be realistically represented in the a priori fluxes.

Appendix

The exponentially decaying temporal autocorrelations is a feature newly implemented into the Jena Inversion System. Temporal correlations are not directly defined as off-diagonal elements in the a-priori error covariance, as the latter does not appear explicitly in the inversion. Rather, the inversion system involves time series filtering in terms of weighted Fourier expansions. More specifically the columns of matrix G_{tcor} contain Fourier modes, weighted according to the frequency spectrum that corresponds to the desired autocorrelation function. The reader is referred to Rödenbeck (2005) for more information. Following Rödenbeck (2005) we define the following spectral weight w:

$$w = \frac{\nu_{low}}{\sqrt{\nu_{low}^2 + (2\pi\nu)^2}} \qquad \textbf{A1}$$

where ν_{low} is the characteristic frequency. The characteristic frequency ν_{low} can be calculated from the desired temporal autocorrelation time (30 days) of the exponential decay and is expressed in years:

ν_{low} = 1/(1/12) where 1/12 is the autocorrelation time in years. Hence the characteristic frequency corresponding to a monthly autocorrelation is 12.

To test numerically whether the implemented autocorrelation decay shape approximates an exponential decay, an error realization of the characteristic frequency was added to the prior fluxes, and the autocorrelation function as described in Kountouris et al. (2015) was calculated numerically simultaneously for the flux time series of all grid cells. Then an exponentially decaying function was fitted (Figure A1) to derive the autocorrelation scale for the corresponding frequency. The resulting autocorrelation shape indeed approximates very well an exponential decay, with an e-folding time of precisely 30 days. The tight confidence bounds of the fitted parameter (29.3 and 30.6 days within 95 % confidence interval), in combination with the small residual sum-of-squares (0.14) suggests a very good approximation of the exponential decay.

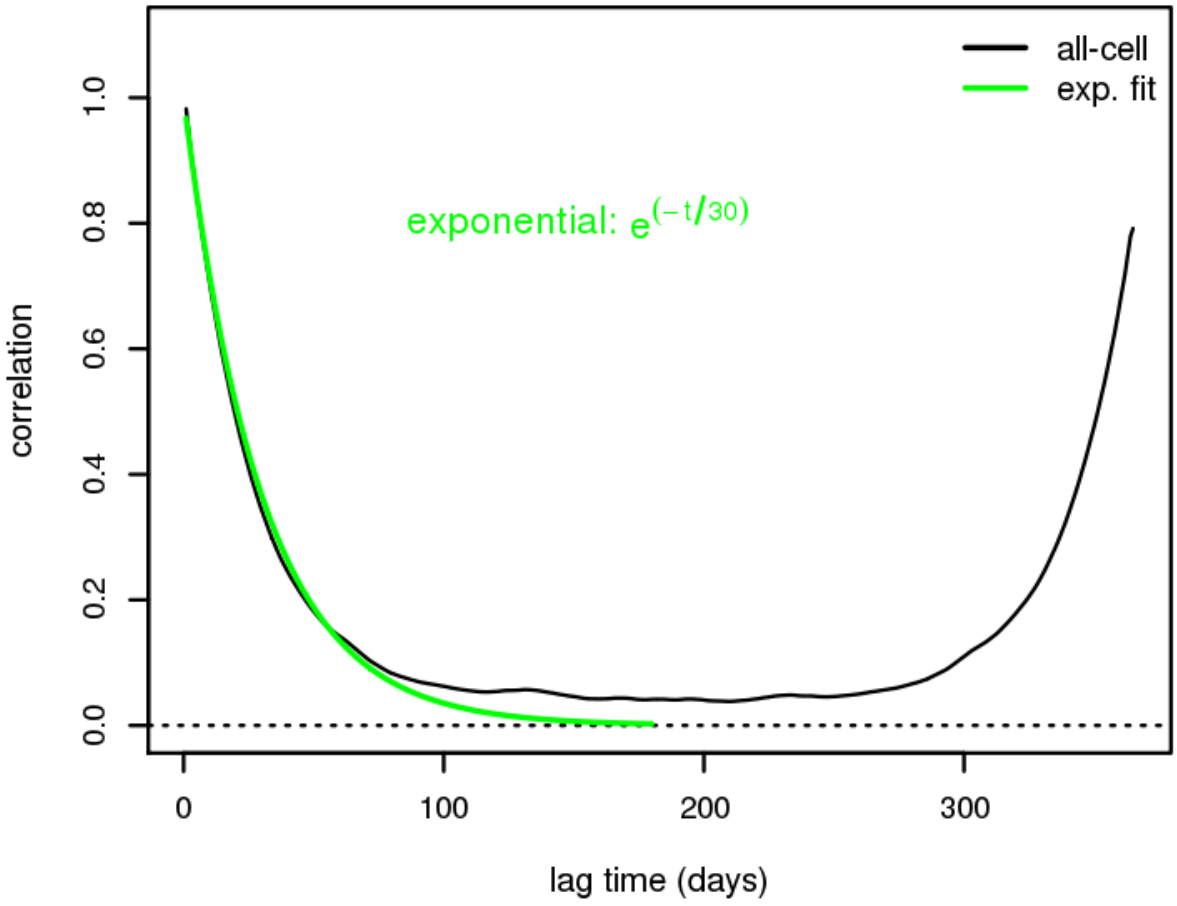

Figure A1 Autocorrelation function for a characteristic frequency of the exponential filter. The autocorrelation is calculated simultaneously for all the domain grid cells. The numerical realization of the autocorrelation does not decay to zero because of the flux seasonality.

Chapter 5

Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part2: the European terrestrial CO_2 fluxes

Abstract

Optimized biogenic carbon fluxes for Europe were estimated from high resolution regional scale inversions, utilizing atmospheric CO_2 measurements at 16 stations. Additional sensitivity tests with different data-driven error structures were performed. As the atmospheric network is rather sparse and consequently contains large spatial gaps, we use a priori biospheric fluxes to further constrain the inversions. The biospheric fluxes were simulated by the Vegetation Photosynthesis and Respiration Model (VPRM) at a resolution of 0.1º and optimized against Eddy covariance data. Overall we estimate an a priori uncertainty of 0.54 GtC y^{-1} related to the poor spatial representation between the biospheric model and the ecosystem sites. The sink estimated from the atmospheric inversions for the area of Europe (as represented in the model domain) ranges between 0.23 and 0.38 GtC y^{-1} (0.30 and 0.49 GtC y^{-1} up-scaled to geographical Europe). This is within the range of posterior flux uncertainty estimates of previous studies using ground based observations.

5.1 Introduction

Global and regional atmospheric inversions have been applied using atmospheric CO_2 measurements made by a global network since two decades to infer terrestrial carbon fluxes using surface in situ or flask measurements of CO_2 dry mole fractions. The optimization of CO_2 biospheric fluxes for the European domain has been the focus of interest in previous studies either using pseudo or real data (Gurney et al., 2004; Peters et al., 2010; Carouge et al., 2010a; Carouge et al., 2010b; Rivier et al., 2010; Broquet et al., 2011; Broquet et al., 2013; Peylin et al., 2013). Retrieved fluxes from most of the inversions are obtained from global systems at coarse resolution which makes difficult to retrieve the spatial and temporal flux variability at finer scales. Large uncertainties in the flux retrievals are introduced due to the coarse resolution of the transport models used and due to the network sparseness (Peters et al., 2010). For example the prevailing westerly winds and the fact that all atmospheric sites are mainly located in central Europe, introduce large flux uncertainties at eastern European regions.

Apart from ground based observations, satellite measurements have also been recently used in atmospheric inversions to infer terrestrial fluxes (Basu et al., 2013; Deng et al., 2014; Chevallier et al., 2014). The advantage of using space-borne measurements lies on the high density of the observations providing the opportunity to constrain regions which the ground network does not see. However satellite based inversions significantly differ from ground based inversions, reporting a larger annual uptake for Europe. A characteristic example is the estimated European uptake in the study by Reuter et al. (2014). They calculated an uptake of 1.02 $GtCy^{-1}$ which

triggered an ongoing debate on whether those estimates are data driven or they lack robustness due to deficiencies in the satellite observations and in the inverse modeling (Feng et al., 2016).

One of the largest sources of uncertainty in inversions is the atmospheric transport uncertainty. Modeled tracer dry mole fractions are biased particularly due to uncertainties of vertical mixing near the surface (Gurney et al., 2003; Gerbig et al., 2008; Houweling et al., 2010). As a consequence, posterior flux estimates are also biased because biases in concentrations are translated into biases in fluxes through the optimization procedure. Propagation of uncertainties in winds (Lin and Gerbig, 2005) and in mixing heights (Gerbig et al., 2008) for summer months with active vegetation resulted in uncertainties in simulated dry air mole fractions of 5.9 ppm and 3.5 ppm respectively.

The current study is a continuation of the Kountouris et al. (2016a) study (hereafter referred to as Ko16) in which the inversion system and its set-up were assessed based on pseudo data. As a next step we apply the modeling framework to real CO_2 atmospheric observations. Our main objectives in the second part of this work are to investigate the potential to infer flux estimates for Europe with reduced uncertainties, and to estimate biospheric fluxes at high spatial resolution and for a full year. We use a spatial flux resolution of 0.25^o x 0.25^o, and the state space allows optimizing 3-hourly NEE corrections to the prior NEE fluxes at a nominal spatial resolution of 0.5^o x 0.5^o. A data driven error structure is implemented consistent with model-data flux mismatches (Kountouris et al., 2015) as tested in part 1 (Ko16) of this study. Further, different error structures are used and assessed including also a spatial error structure with a hyperbolic correlation shape as suggested by Chevallier et al. (2012). Since spatial autocorrelations have been found to be very short, the annual aggregated uncertainty over the European domain is smaller than traditionally assumed (see also Ko16). The error inflation necessity and implementation was addressed in Ko16 either by inflating the error covariance or, more formally, by introducing a bias term. However the hyperbolic correlation shape suggested by Chevallier et al. (2012) has a stronger impact from larger distances compared to the exponential shape, leading to an aggregated uncertainty which does not require to be inflated. We perform also a number of sensitivity tests to account for misrepresentation of the fossil fuel signal and also for transport uncertainties due to vertical mixing.

This chapter is structured as follows: Section 5.2 describes the inversion system, the network and station data which are used and details the assumed error structure. Section 5.3 shows the results of the goodness of fit, and the retrieved fluxes. The data fitting and the reliability of the posterior fluxes are extensively discussed in section 5.4.

5.2 Methods

5.2.1 Two-step inversion

Real-data inversions require a nested inversion scheme, since observations contain also contributions from regions outside of the Domain of Interest (DoI). As in part 1 of this study (Ko16), the Jena Inversion System (Rödenbeck 2005) including the two-step nesting scheme (Rödenbeck et al., 2009; Trusilova et al., 2010) was used. This scheme allows for combining regional and global inversions within a consistent system. Here we only provide a brief description as details are given in Rödenbeck et al. (2009) and Trusilova et al. (2010). The

atmospheric transport models TM3 (step 1) (Heimann and Körner, 2003) and STILT (step 2) (Lin et al., 2003) were used for transport at the global and regional domain, respectively. For the global runs, TM3 was used at a spatial resolution of 4^o latitude x 5^o longitude, driven by meteorological fields from the ERA-Interim reanalysis produced by ECMWF (Dee et al., 2011). The transport matrix for the regional inversions, was identical to the one used for the synthetic data study in part 1.

In the first step, a global inversion is performed using the global transport model. The outcome is an optimized flux field, at coarser scale for the full period (FP) and the global domain. Then two forward runs are performed. The first run uses the global transport model over the FP, computing the modeled mixing ratios Δc_{mod1}. The second run initializes again the global transport model but only within the regional DoI. This can be regarded as a regional simulation, but with coarse resolution, yielding modeled mixing ratios Δc_{mod2}. Then the "remaining mixing ratio" is calculated for all the observing sites inside the DoI:

$$\Delta c_{remain} = c_{meas} - (\Delta c_{\mathrm{mod1}} - \Delta c_{\mathrm{mod2}} + c_{ini}) \qquad \textbf{5.1}$$

were c_{ini} the initial condition which corresponds to a well mixed atmosphere with a given initial tracer mixing ratio.

In step two, the high-resolution transport model is used for the regional inversion within the DoI, where all fluxes are represented at fine resolution. For this inversion the vector containing the measured mixing ratios c_{meas} are replaced by the "remaining mixing ratios" Δc_{remain}. The optimized fluxes from this step are the high-resolution fluxes of interest.

5.2.2 Atmospheric network and data

Table 5.1 Information on the stations used for the regional inversions. Same network applied for the synthetic, and the real data inversions in Kountouris et al. (2016). In first column the term "type" stands for continuous (C) or flask (F) data. Under "Data origin" WDCGG means "World Data Centre for Greenhouse Gases".

Site Code / type	Name	Latitude (^{o})	Longitude (^{o})	Height (m.a.s.l.) (m)	Measurement height (above ground) (m)	Model height	Data provider	Data origin	Citation
BAL/F	Baltic Sea, Poland	55.50	16.67	8	57	28	NOAA	Direct contact	Dlugokencky et al. 2015
BIK/C	Bialystok, Poland	53.23	23.03	183	90	90	MPI-BGC	Direct access	Popa et al. (2010)
CBW/C	Cabauw, Netherlands	51.58	4.55	-2	200	200	ECN	Direct contact	Vermeulen et al. (2011)
CMN/C	Monte Cimone, Italy	44.18	10.7	2165	12	670	IAFMS	WDCGG	Alemanno et al. (2014)
HEI/C	Heidelberg, Germany	49.42	8.67	116	30	30	University of Heidelberg	CarboEurope	Hammer et al. (2008)
HPB/F	Hohenpeissenberg, Germany	47.80	11.01	934	50	10	NOAA	Direct contact	-
HUN/C	Hegyhatsal, Hungary	46.95	16.65	248	115	96	HMS	WDCGG	Haszpra et al. (2001)
JFJ/C	Jungfraujoch, Switzerland	46.55	7.98	3572	10	720	University of Bern	CarboEurope	-
KAS/C	Kasprowy Wierch	49.23	19.93	1987	5	480	UKRAK, AGH	CarboEurope	Necki et al. (2013)

LMU/C	La Muela, Spain	41.36	-1.6	570	79	80	University of Barcelona	CarboEurope	-
MHD/C	Mace Head, Ireland	53.33	-9.90	25	10	15	LSCE	WDCGG	Ramonet et al. (2010)
OXK/C	Ochsenkopf, Germany	50.03	11.81	1022	163	163	MPI-BGC	CarboEurope	Thompson et al. (2009)
PRS/C	Plateau Rosa, Italy	45.93	7.71	3480	-	500	RSE	WDCGG	Ferrarese et al. (2015)
PUY/C	Puy De Dome, France	45.77	2.97	1465	10	400	LSCE	CarboEurope	Lopez et al. (2015)
SCH/C	Schauinsland, Germany	47.92	7.92	1205	-	230	UBA	WDCGG	-
WES/C	Westerland, Germany	54.93	8.32	12	-	15	UBA	WDCGG	-

Glossary for the data providers: AGH: University of science and Technology Polland, ECN: Energy research Centre of the Netherlands, HMS: Hungarian Meteorological Service, IAFMS: Italian Air Force Meteorological Service, LSCE: Le Laboratoire des Sciences du Climat et de l'Environnement, MPI-BGC: Max Planck Institute for BioGeoChemistry, NOAA: National Oceanic and Atmospheric Administration, RSE: Ricerca sul Sistema Energetico, UBA: Umweltbundesamt, UKRAK: Department of Environmental Physics Polland

For step 1 we used the same station network as in version s04_v3.6 of the Jena Carboscope CO_2 inversion (http://www.bgc-jena.mpg.de/CarboScope/?ID=s04_v3.6), with 64 stations globally. For step 2 (regional inversion) continuous and flask measurements from 16 stations within Europe were used as described in Ko16 (see also Table 5.1). Of those 16 stations 7 are already included in the step 1 inversion. All provided valid values were used, except those paired flask measurements that differ more than 0.34 ppm which were omitted. Measurements from the continuous stations were aggregated to hourly values where needed, and neighboring values differing by more than 1 ppm were omitted. Night and day time observations were selected depending on the type of station (Ko16). As all institutions report mixing ratio values traceable to WMO (World Meteorological Organization) calibration scale, we expect compatibility between the different datasets (also see Rödenbeck et al., 2006).

Table 5.2 Overview of the inversion scenarios. "Shape" describes the internal structure of the bias component (proportional to respiration R or to Net Ecosystem Exchange NEE), and "Time vary" indicates whether the bias component also has temporal variations or not. The fifth column "Prior" represents the terrestrial model used as prior, and "Correlation shape" describes the functional form used for the spatial prior uncertainty correlation, either exponential (E) or hyperbolic (H). The last column indicates whether the full or the reduced station network was assumed.

Inversion code	Bias component	Shape	Time vary	Prior	Correlation shape	No. of Stations
B1	-	-	-	VPRM	E	16
B2	-	-	-	GBIOME	E	16
S1	Yes	R	Flat	VPRM	E	16
S1a	Yes	NEE	Flat	VPRM	E	16
S1b	Yes	R	Vary	VPRM	E	16
S2	-	-	-	VPRM	E	14
S3	-	-	-	VPRM	H	16

In this study we use the site HEI (Heidelberg) which is traditionally not used for European CO_2 flux inversions as being considered too local (Broquet et al., 2013; Rödenbeck et al., 2009; Rivier et al., 2010). The Heidelberg region is considered to be one of the most polluted regions in Germany (Fiedler et al. 2005) and therefore could bias the flux estimates. Moreover the WES (Westerland) site contains long periods with no data. This could potentially affect posterior flux estimates since extended data gaps can lead to jumps in the presence of biases. Thus we evaluate the performance and the sensitivity of the European flux estimates to the network configuration, by performing also an inversion (referred to as S2, see

Table 5.2) excluding HEI and WES.

5.2.3 A-priori information and uncertainties

A set of inversion cases differing in the prior information, the error structure and the station configuration was realized (see overview in

Table 5.2). Prior information derived from both biosphere models (VPRM and GBIOME-BGCv1) is used to investigate the impact of the prior fields to the posterior flux estimates. Furthermore an ensemble of inversions using different error structures is used to investigate the impact on the posterior flux estimates and uncertainties.

Similarly to the synthetic inversion (Ko16) we use the base case B1 which inflates the prior uncertainty by up-scaling the error covariance matrix, and case S1 which includes a bias term. In the base case the VPRM model provides the prior flux field, and exponentially decaying correlations are assumed. The bias component in the S1 scenario will always have a correction with the same sign for all grid-cells as it just scales a predefined flux field. In the S1 case it follows the shape of the annually averaged respiration flux, in the S1a case that of the a priori net biogenic flux, and in the S1b case again that of the annually averaged respiration flux, but with monthly temporal resolution of the bias term to allow for some temporal flexibility. The B2 inversion refers to the scenario where GBIOME-BGCv1 was used as a priori information instead of VPRM, and the error structure does not contain a bias term. With this case we can evaluate how sensitive the posterior flux estimates are with respect to the prior information which has been used. We also examine a spatial error structure based on a hyperbolic (instead of an exponential) spatial correlation shape as suggested in Chevallier et al. (2012) which we will refer to as S3 scenario.

Note that in most of the inversions performed, VPRM fluxes were used as prior information. Those fluxes are already optimized using EC measurements, therefore evaluation of the posterior flux estimates against EC data at the local scale could result in posterior fluxes that are limited or even not further constrained (since they are already optimized). In contrast, posterior fluxes produced with BIOME-BGC used as prior are expected to show significantly larger corrections compared to the prior estimations, and are therefore used for evaluation against EC data. Nevertheless in most cases we use VPRM as prior in order to keep our estimates as data-driven as possible through the overall optimization procedure; at local scale by using EC data, and at regional scale using the atmospheric dry mole fractions.

As in the synthetic experiment (Ko16) the temporal decorrelation time was set to 31 days. In Kountouris et al. (2015), model-data comparisons representative at site scale (around 1 km) showed spatial correlation lengths of 40 km whilst model-model comparisons representative at 50 km resolution identified a correlation scale of 370 km. Considering also that the state space has a resolution of 50 km, the spatial decorrelation length was chosen to be approximately 100 km (66 km in meridional, and 130 km in zonal direction). In the prior error covariance, diagonal elements of 2.27 $\mu molm^{-2}s^{-1}$ were assumed, consistent with the model-data flux mismatches as calculated in Kountouris et al. (2015). Propagating this spatiotemporal error structure yields a domain-integrated uncertainty (E_{st}) of 0.15 GtC y^{-1}. Note that this is substantially smaller than for the synthetic experiment due to the much shorter spatial correlation length scales. A total annual, domain integrated uncertainty E_{tot} of 0.3 GtC y^{-1} was assumed, which corresponds to twice the standard deviation of annual terrestrial flux estimates for 2007 between terrestrial biosphere models taken from the global carbon atlas (http://www.globalcarbonatlas.org). This is also consistent with the prior uncertainty (for Europe) assumed for the global inversions

performed by the Jena inversion system. For those inversions in which the additional bias term was considered, its error E_{BT} was calculated using

$$E_{tot}^2 = E_{ST}^2 + E_{BT}^2$$

5.2

For the S3 scenario using hyperbolic correlations similar to Chevallier et al. (2012) ($\frac{1}{1+d}$), a characteristic value d (lag distance) was used such that the correlation drops after around 60 km to 1/e of its initial value, consistent with the hyperbolic fit to the model-data flux residual autocorrelation in Kountouris et al. (2015). For this case no additional bias term was needed, as the spatially and temporally aggregated uncertainty was found to be 0.32 GtC y^{-1}, which is very close to the uncertainty assumed for the inversions (0.3 GtC y^{-1}).

Furthermore, we include ocean fluxes from Mikaloff-Fletcher et al. (2007), and anthropogenic emissions from the EDGAR v4.1 inventory scaled at national level for individual years according to the BP (British Petroleum) statistical review of world energy (BP, 2012) following Steinbach et al. (2011). Anthropogenic emissions are considered to be perfectly known (with no prior uncertainty), as one typically assumes that there is more a-priori knowledge regarding the anthropogenic emissions as compared to biogenic fluxes. As the inversion cannot distinguish between biogenic and anthropogenic signals, any errors in the a-priori anthropogenic emissions will be included as corrections to the NEE flux.

5.2.4 Diagnostics and aggregation of fluxes

Similar to Ko16 we use the χ_c^2 metric to evaluate the goodness of fit for each station (Eq. 5.3)

$$\chi_c^2 = \frac{\sum_t \frac{(\Delta c_t)^2}{\sigma_t^2}}{n}$$

5.3

where Δc_t is the model-data mismatch in dry mole fractions for a given observation time t, n the number of observations and σ_t the assumed uncertainty. Further we make use also of the reduced χ_r^2 (Eq. 5.4) where J_{min} is the cost function at its minimum

$$\chi_r^2 = 2\frac{J_{\min}}{n}$$

5.4

For more details about the chi-square metric the reader is referred to Ko16 study.

The optimized fluxes are derived at 0.25º spatial and daily temporal resolution from the inversion system. We post-process the fluxes by aggregating them spatially at country/domain-wide scales and temporally at monthly/annual scales.

Flux comparisons with other studies require that both fluxes refer to the same geographical region. Typically studies refer to TransCom regions with a European domain that expands more

into the Eurasian region. To scale our results to the TransCom EU region, we calculated the flux ratio between the TransCom EU region and our European domain from the first step of the two-step inversion (global inversion), and used this ratio (about 1.3) to scale our posterior estimates and the corresponding uncertainties (presented in Figure 5.7).

5.3 Results

5.3.1 Simulated CO_2 and goodness of fit

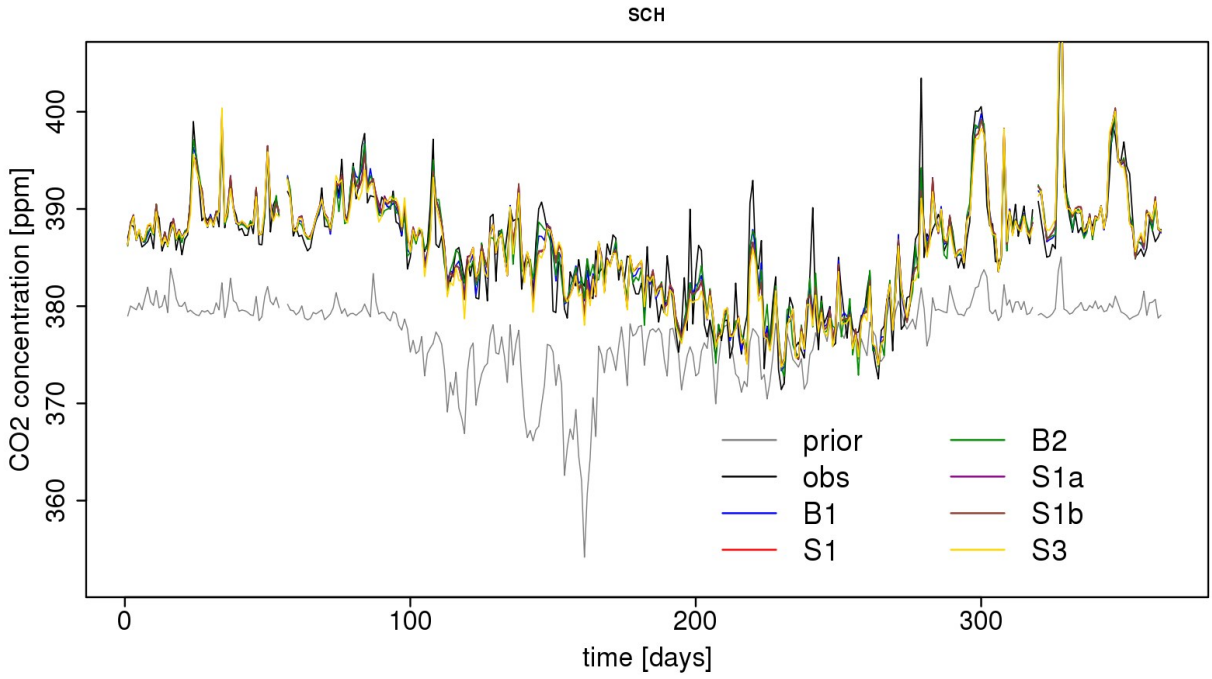

Figure 5.1 Daily nighttime (23:00-4:00 UTC) averages for prior, true, and posterior CO_2 dry mole fraction time series for the Schauinsland site for the real data inversion. Time starts at 1st January 2007.

Figure 5.1 presents a comparison of observed and modeled daily averages of the nighttime (hours 23, 00, 1, 2, 3, 4 UTC) CO_2 dry air mole fractions for the Schauinsland station (SCH), a mountain station, for the year 2007. The prior estimates (gray line) as derived from a forward model run using VPRM flux fields are systematically lower than the observations (black line) with the most divergent values occurring during the growing season. A similar pattern was found for the other atmospheric stations. Posterior CO_2 timeseries from all the inversions are in much closer agreement with the observations.

Table 5.3 RMSD (first column in ppm) and correlation coefficients (second column) between observations and prior/posterior CO_2 dry mole fractions for daily "daytime" or "nighttime" averaged values and for each station. The third column shows χ_c^2, the normalized dry mole fraction mismatch per degree of freedom for 7-day averaged residuals, as a measure of how well the data were fitted. The format for each station is as follows: RMSD | r^2 | χ^2.

	Prior	B1	B2	S1	S1a	S1b	S2	S3
BAL	7.12 | 0.20 | 69.35	1.48 | 0.97 | 0.89	1.53 | 0.97 | 0.93	2.26 | 0.93 | 2.04	2.26 | 0.93 | 2.03	2.25 | 0.93 | 2.02	1.41 | 0.97 | 0.83	2.37 | 0.92 | 2.07
BIK	8.20 | 0.52 | 60.10	2.93 | 0.93 | 0.88	3.17 | 0.92 | 0.99	3.52 | 0.90 | 1.51	3.52 | 0.90 | 1.53	3.51 | 0.90 | 1.53	2.93 | 0.93 | 0.88	3.78 | 0.88 | 1.70
CBW	8.71 | 0.23 | 83.98	3.43 | 0.88 | 2.05	3.49 | 0.88 | 2.18	4.09 | 0.83 | 2.47	4.09 | 0.83 | 2.48	4.09 | 0.83 | 2.49	3.42 | 0.88 | 1.99	4.33 | 0.81 | 2.61
CMN	4.20 | 0.40 | 31.73	1.26 | 0.96 | 0.16	1.35 | 0.95 | 0.19	1.45 | 0.94 | 0.19	1.44 | 0.95 | 0.19	1.46 | 0.94 | 0.21	1.25 | 0.92 | 0.15	1.57 | 0.94 | 0.26
HEI	14.04 | 0.37 | 31.28	6.93 | 0.84 | 3.05	7.07 | 0.83 | 3.07	7.92 | 0.79 | 4.22	7.91 | 0.79 | 4.23	7.92 | 0.79 | 4.23	-	8.34 | 0.77 | 5.17
HPB	5.06 | 0.43 | 15.61	1.41 | 0.91 | 0.34	1.70 | 0.94 | 0.50	2.00 | 0.96 | 0.65	2.01 | 0.91 | 0.66	2.00 | 0.91 | 0.65	1.41 | 0.96 | 0.33	2.03 | 0.91 | 0.67
HUN	7.44 | 0.55 | 66.36	2.58 | 0.94 | 0.84	2.74 | 0.93 | 0.88	3.07 | 0.92 | 1.32	3.08 | 0.92 | 1.34	3.08 | 0.92 | 1.33	2.58 | 0.94 | 0.87	3.43 | 0.90 | 1.98
JFJ	4.52 | 0.03 | 21.39	1.96 | 0.77 | 1.59	2.23 | 0.72 | 1.53	2.07 | 0.75 | 1.83	2.07 | 0.75 | 1.82	2.07 | 0.75 | 1.84	1.95 | 0.78 | 1.58	2.10 | 0.74 | 1.98
KAS	6.35 | 0.39 | 52.58	3.41 | 0.87 | 2.90	3.43 | 0.87 | 2.89	3.88 | 0.82 | 3.96	3.88 | 0.82 | 3.99	3.87 | 0.83 |3.93	3.29 | 0.77 | 2.77	4.01 | 0.81 | 4.67
LMU	6.01 | 0.05 | 29.00	1.45 | 0.94 | 0.29	1.51 | 0.94 | 0.28	1.74 | 0.92 | 0.59	1.74 | 0.92 | 0.58	1.76 | 0.92 | 0.60	1.44 | 0.95 | 0.29	1.84 | 0.91 | 0.68
MHD	4.50 | 0.21 | 22.24	1.23 | 0.94 | 0.24	1.20 | 0.94 | 0.21	1.29 | 0.92 | 0.31	1.74 | 0.93 | 0.31	1.76 | 0.94 | 0.31	1.23 | 0.94 | 0.24	1.26 | 0.94 | 0.27
OXK	5.39 | 0.28 | 38.95	2.45 | 0.85 | 0.79	2.52 | 0.84 | 0.85	2.78 | 0.81 | 1.19	2.78 | 0.81 | 1.20	2.79 | 0.81 | 1.20	2.41 | 0.86 | 0.70	2.98 | 0.78 | 1.59
PTR	2.98 | 0.07 | 20.75	1.06 | 0.89 | 0.46	1.10 | 0.88 | 0.49	1.16 | 0.87 | 0.52	1.16 | 0.87 | 0.52	1.17 | 0.87 | 0.52	1.07 | 0.89 | 0.45	1.22 | 0.86 | 0.53

PUY	4.86 | 0.29 | 39.48	2.05 | 0.87 | 0.67	2.16 | 0.86 | 0.75	2.40 | 0.82 | 0.97	2.40 | 0.82 | 0.97	2.40 | 0.82 | 0.95	2.02 | 0.88 | 0.71	2.48 | 0.81 | 1.27
SCH	5.18 | 0.24 | 41.77	1.90 | 0.89 | 0.27	2.00 | 0.88 | 0.28	2.23 | 0.85 | 0.51	2.23 | 0.85 | 0.51	2.23 | 0.85 | 0.51	1.84 | 0.90 | 0.24	2.38 | 0.84 | 0.70
WES	8.06 | 0.23 | 41.77	2.21 | 0.94 | 0.27	2.00 | 0.94 | 0.28	2.23 | 0.91 | 0.51	2.23 | 0.91 | 0.51	2.23 | 0.91 | 0.51	-	2.38 | 0.90 | 0.70

Table 5.3 summarizes the statistics between the modeled and the observed CO_2 dry mole fractions for all stations based on daily averages using the respective sampling times for mountain and other stations. Of note is that the real data inversions include errors due to the modeling of transport, which is not the case in the synthetic experiment in Ko16 as the same transport model was used for forward and inversion runs. Standard deviations of the posterior residuals (observed – modeled) show an average decrease for all inversion setups and for all stations of 59% compared to the prior residuals. Correlations between prior and observed as well as posterior and observed mole fractions (also Table 5.3) were likewise increased on average from 0.48 to 0.93. Of note is that B1 and B2, which use an inflated prior error covariance for the spatiotemporal component, show larger improvement relative to the prior in RMSD and some limited improvement in correlation coefficient, compared to those inversions where a bias component was included (S1, S1a, S1b). Figure 5.2 visually summarizes the goodness of fit in a Taylor diagram for cases B1 and S1, presenting prior and posterior estimates of the correlation and the normalized standard deviation between modeled and observed CO_2 dry mole fraction time-series. It is obvious that the additional flexibility of B1 in the spatiotemporal flux distribution results in a better reproduction of the concentration variability. The same picture emerges when comparing the B1 and B2 inversions to S3 (see Table 5.3). Although all these cases assume no explicit bias term in the error structure, the larger correlations from areas farther away for the S3 case with a hyperbolic correlation causes a reduced number of effective degrees of freedom, which results in larger residuals in posterior-observed mole fractions (Table 5.3) comparable to those of the S1 case.

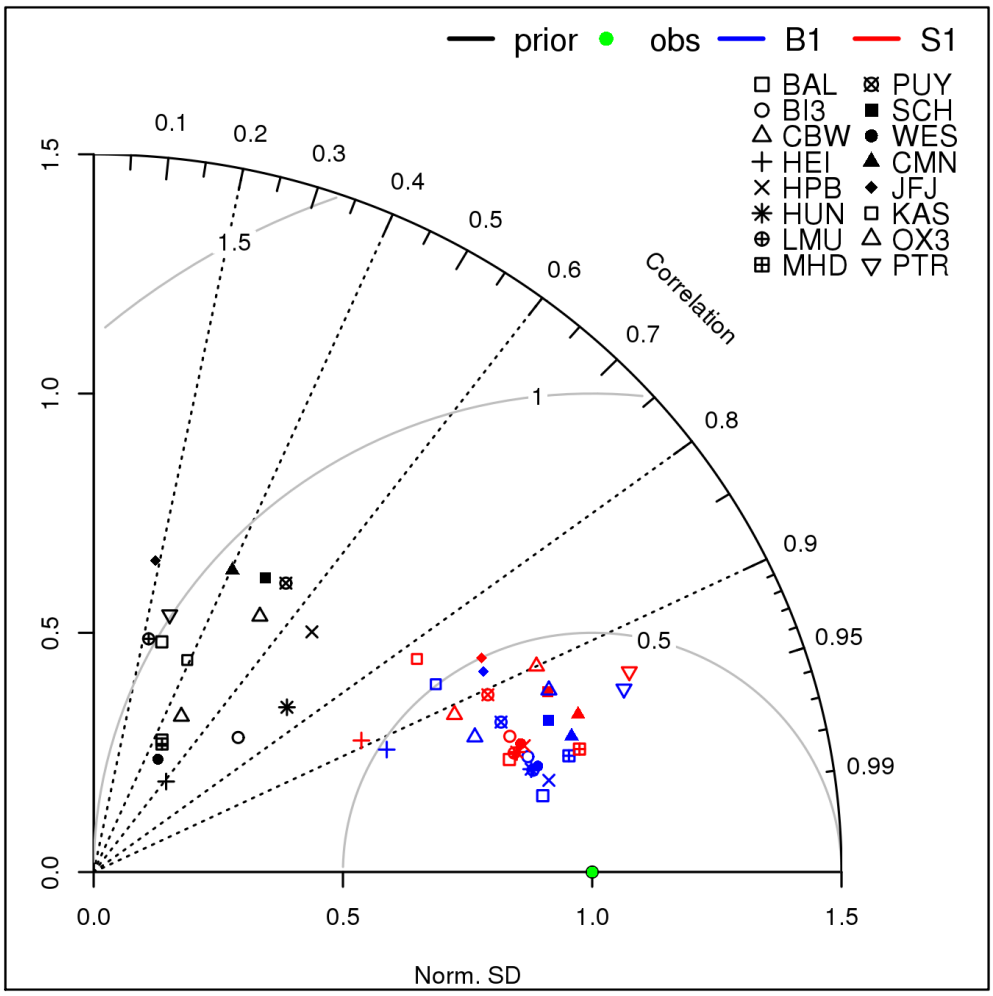

Figure 5.2 Taylor diagram for modeled and observed time-series of CO_2 dry mole fractions. Prior (black), observed (green, the perfect match of modeled and observed time-series) and the different inversion cases (B1 blue; S1 red) are displayed. Different symbols denote different atmospheric stations. The normalized SD was calculated as the ration of the SD of the modeled time-series to the SD of observations. Gray semi-circles show contours of the standard deviation of the model error.

Calculating the goodness of fit using the station-specific χ_c^2 values from Eq. 5.3, most of the sites (Table 5.3) show values around 1, indicating that the misfits are inside the uncertainty range. For the CBW, HEI, JFJ, KAS sites, values above 1 regardless the error structure were found, with the most extreme value of 5.17 for the HEI site in the S3 inversion. This could suggest that for a polluted site as HEI larger uncertainties should be considered.

The reduced χ_r^2 values regarding the overall model performance (Eq. 5.4) for all inversion set ups is found to be close to 1 with χ^2 values of 1.08 (B1), 1.16 (B2), 1.17 (S1), 1.17 (S1a), 1.19 (S1b), 0.89 (S2) and 1.25 (S3), suggesting that the assumed prior uncertainty describes well the actual uncertainties.

5.3.2 Posterior flux estimates at different scales

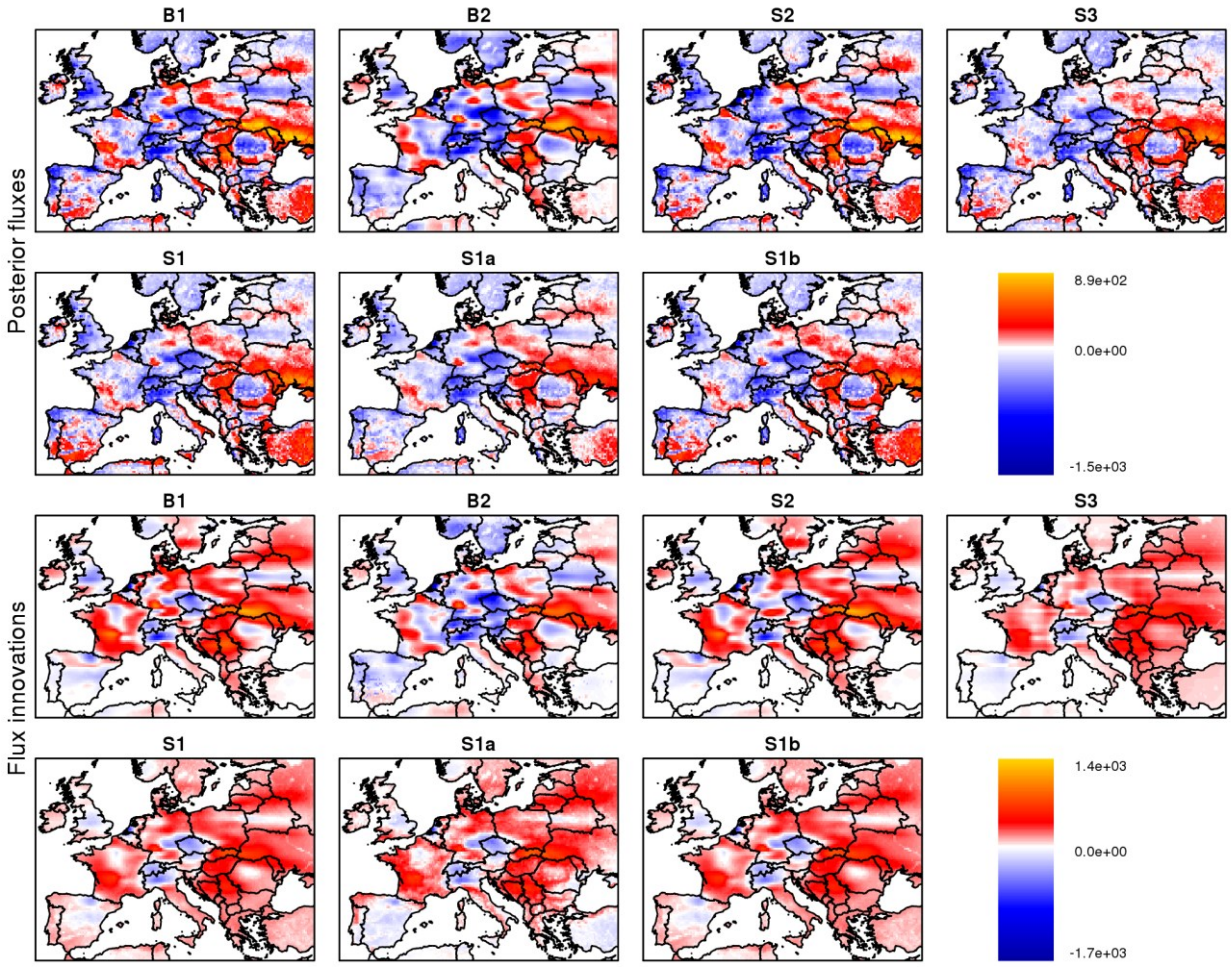

Figure 5.3 Annual biogenic flux spatial distribution (top two rows) and flux innovations (posterior - prior) (bottom two rows) as estimated from the different inversions for the real data case. Units are in $gCy^{-1}m^{-2}$.

The annually integrated spatial flux distribution is presented in Figure 5.3 for all the different inversion settings. Differences between the results based on the two general error structures (with and without the bias term) were observed mainly in central and Western Europe (longitudes less than 20º E), where the network provides a strong constraint. This difference is characterized by stronger spatial flux variability for the general B1 case, with multiple transitions between carbon sources and sinks at regional scales. The same picture emerges for the western part of Europe. In contrast, all the inversions including a bias component (S1, S1a, S1b) yield a more homogeneous flux distribution with somewhat finer structure in the flux retrievals (e.g France and north-east part of Europe). Comparisons between S1, S1a, S1b flux distributions do not show any significant difference. Almost the same picture emerges when comparing B1 and S2 cases, indicating that excluding the 2 stations does not have a very strong influence on our annual flux estimates. However spatial differences were observed for the areas close to the two sites. The most important one applies for the area near the HEI station where we observed a transition from source to net carbon sink when excluding the corresponding site. The choice of the prior does only have a small impact on the mean flux as can be seen by comparing posterior fluxes from B1 and B2 despite the significant differences in the flux innovations (Figure 5.3). All innovations show that positive fluxes were added mainly in central Europe and more intensively for the cases

were no bias term was used. Overall the results suggest that the general error structure matters, i.e. whether or not to include a bias term, but how the bias is implemented is of less importance for the retrieved flux patterns. One would expect that the flux distribution from the S3 case would follow the general flux structure from the inversions without the bias term. Interestingly the distribution is similar to the one obtained from the inversions with the bias term (cases S1, S1a, and S1b). This shows that inversions assuming correlations with a strong contribution from the far field have similar characteristics as inversions that assume a flat bias term.

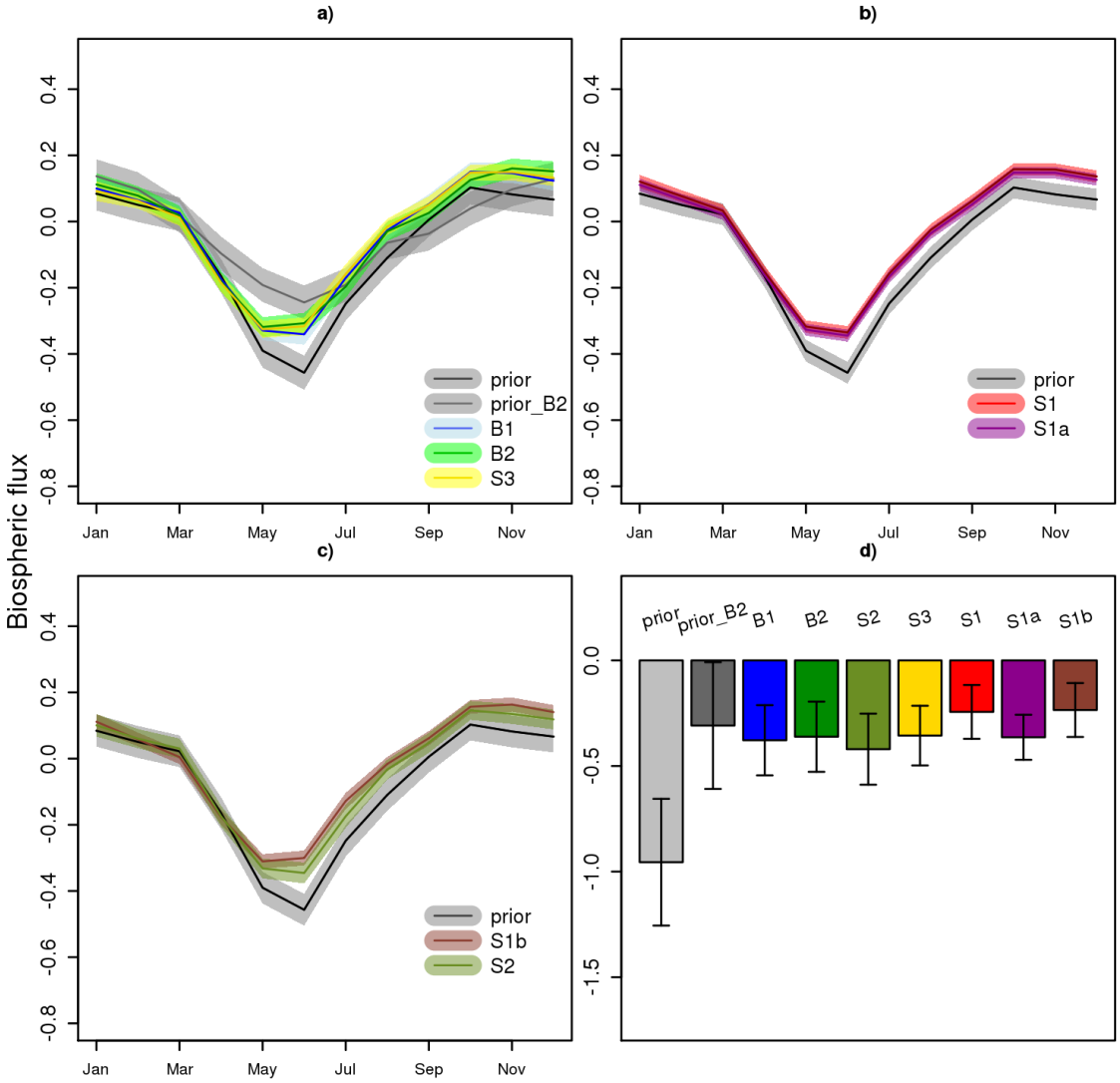

Figure 5.4 Monthly and annual (panel d) biosphere fluxes integrated over the domain. Panel a) shows B1, B2 and S3 cases, b) S1 and S1a and the c) panel shows S1b and S2 cases. Note that all inversions share the same annual prior uncertainty but monthly prior uncertainties differ. Units are in GtC month^{-1} and GtC y^{-1} for monthly and annual fluxes, respectively

Figure 5.4 shows the spatially aggregated posterior flux estimates for the full domain with the corresponding uncertainties integrated at monthly and at annual temporal scales. The same prior uncertainty was used for cases B1 and B2 although they differ in prior flux field. Posterior estimates from B1 (blue line/shading) and B2 (green line/shading) inversions do not show any significant difference at monthly and annual scales despite the large difference in prior fluxes. We observe that the maximum uptake occurs slightly earlier for the B2 case. Monthly fluxes

from the S3 inversion also show the same temporal evolution. We do not observe any significant difference in monthly fluxes for the S1 (red line/shading) and S1a (violet line/shading) inversions. Both cases are comparable to the B1 and B2 cases at monthly and annual scales. A slightly different picture emerges from the S1b inversion, where the bias term allowed for more degrees of freedom for monthly corrections. The resulting seasonal cycle is somewhat smaller, with reduced summer carbon uptake. Inversions that included the bias term yielded smaller posterior uncertainties at both temporal scales, which is expected as the spatiotemporal component of the uncertainty was not inflated as was the case for the B1 scenario. Flux retrievals from the reduced network (sensitivity case S2) show a slightly deeper sink, but the differences to the base case B1 are insignificant (i.e. clearly within the posterior uncertainties).

All of the inversions suggest Europe to be a carbon sink, with a range of -0.23 ± 0.13 GtC y^{-1} to -0.38 ± 0.17 GtC y^{-1} for the S1b and B1 inversions respectively. The mean annual posterior flux estimate for Europe averaged over different inversions amounts to - 0.32 GtC y^{-1}.

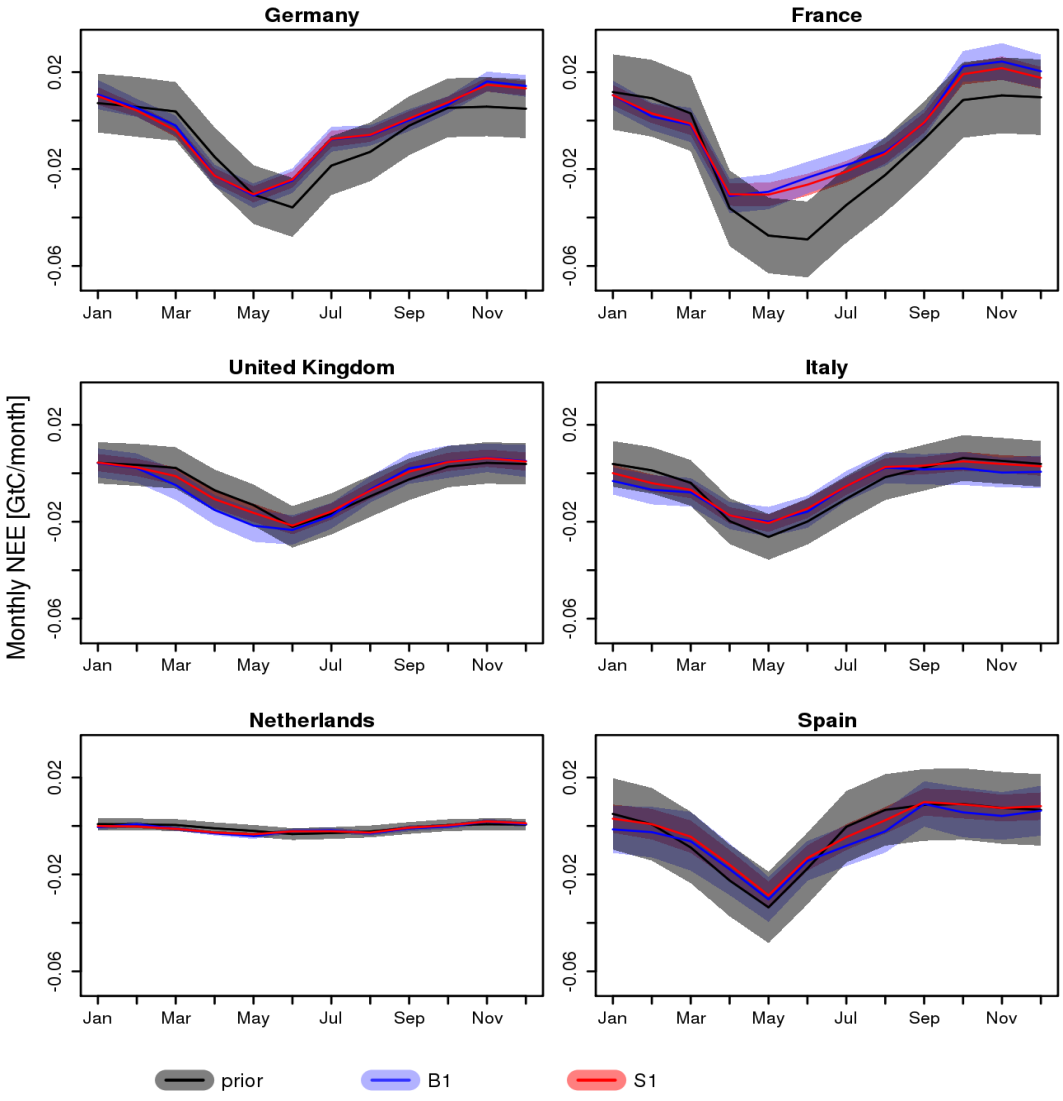

Figure 5.5 Temporal evolution of prior and posterior monthly NEE for selected European countries.

Posterior monthly flux estimates at smaller spatial scales (country level) are shown in Figure 5.5. Areas that are not well constrained by the current network show some divergence in the posterior flux estimates although not significant considering the uncertainty range. For example Germany,

which is better constraint, shows a limited spread of the posterior fluxes with an annually averaged standard deviation between the different posterior flux estimates being 0.0009 GtC y^{-1}, while United Kingdom (which is less well constrained) shows a slightly larger spread of the posterior estimates with an annually averaged standard deviation of 0.002 GtC y^{-1}. Note that the posterior uncertainties are smaller by about 36% for the S1 case, which is related to the smaller prior uncertainties at monthly time scales (see also section 4.3.2).

5.3.3 Validation against eddy covariance measurements

As shown in chapter 4, eddy covariance measurements in principle have the potential for quantitative evaluation of the retrieved fluxes from the inversions. Here we used posterior flux estimates from the B2 inversion for evaluation against eddy covariance measurements, as the prior flux fields in B2 (GBIOME-BGCv1) were not optimized using eddy covariance measurements. Gap-filled data were downloaded from the European Fluxes Database Cluster (http://www.europe-fluxdata.eu). A modified flux-site network compared to the one reported in Kountouris et al. (2015) was used. Specifically we omitted sites that they have not been used for the VPRM optimization (*CH-Fru, CH-Lae, CH-Oe1, ES-LMa, FR-Avi, FR-Mau, IT-Cas, IT-LMa, IT-Ro2, NL-Dij, NL-Lut, SE-Sk1, SK-Tat*) as well as sites that were not available as gap-filled data (*CH-Dav, ES-Agu, FR-Aur*). Further some more sites were added both for the VPRM optimization and for the flux comparisons (*CZ-wet, DK-Sor, HU-Bug, IT-Non, NL-Ca1, PL-wet, RU-Fyo, UK-PL3*). Monthly averaged fluxes were extracted, with weights for each vegetation class that compensate for the asymmetry between number of flux towers per vegetation type and the fraction of land area covered by the specific vegetation type, similar to chapter 4.

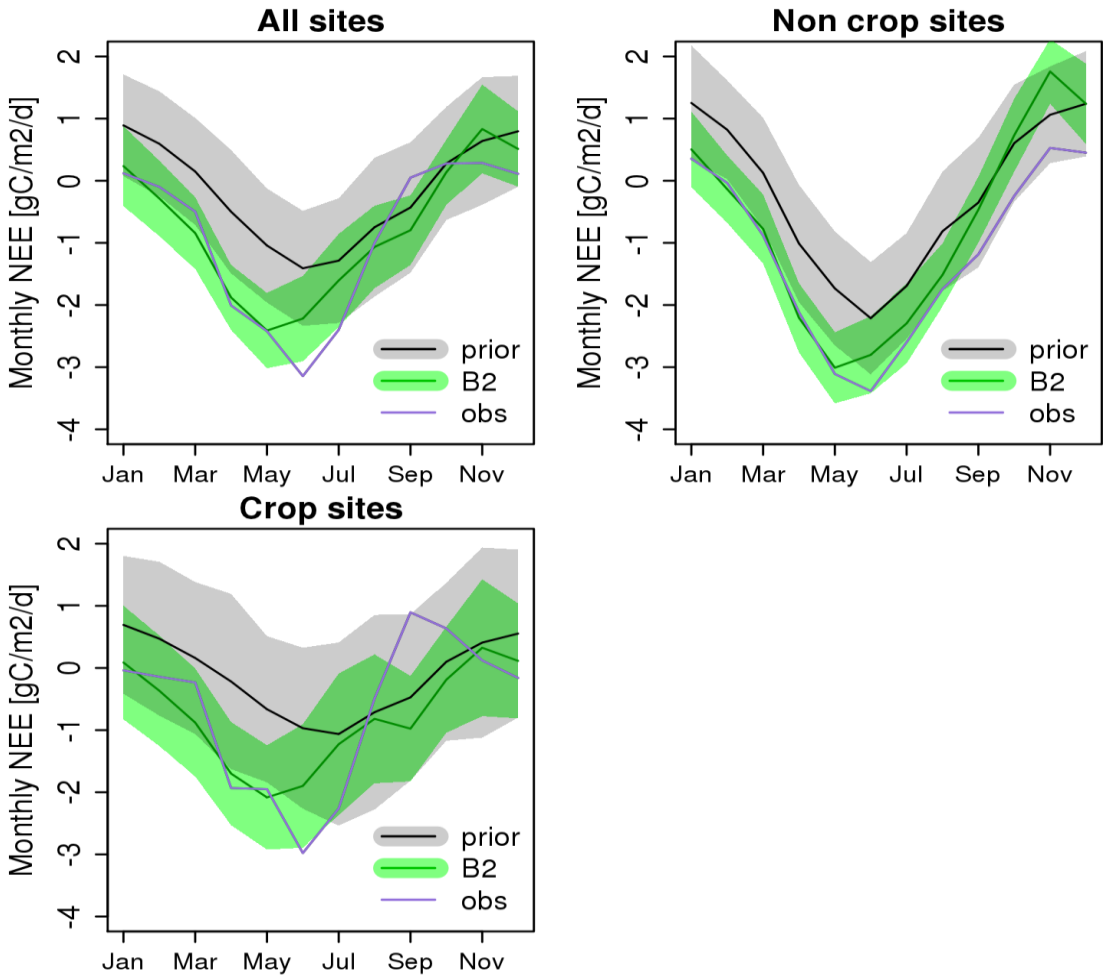

Figure 5.6 Temporal evolution of monthly NEE ($gCm^{-2}day^{-1}$) averaged over all EC sites (top left), excluding crop (top right), and using only crop sites (bottom). Uncertainties (error of the mean monthly NEE) are indicated by the shaded areas.

The analysis of the monthly prior biospheric fluxes in Figure 5.6 reveals significant differences between observed and prior fluxes from the inversion. The GBIOME-BGCv1 model systematically overestimates the observed fluxes throughout the year. The retrieved fluxes from the inversion (dark green line) are closer to the observed fluxes, with a stronger uptake compared to the prior during spring and summer time. The timing of the peak uptake is shifted to one month earlier in comparison to the observations. The mean absolute bias (averaged absolute differences between prior/posterior and observed fluxes) is significantly reduced by 52% from 0.84 to 0.40 gCm^2day^{-1}. The standard deviation of the residuals is reduced by around 24%, from 0.68 for the prior to 0.40 gCm^2day^{-1} for the posterior residuals. Splitting the sites into two main categories, the first only with crops, and the second with non crop sites, revealed differences on how well those sites can be represented. Clearly best matches were found for the non crop sites with a reduction in the mean absolute bias of 51% whilst for the crop sites it is limited to 38%.

5.4 Discussion

We performed a series of atmospheric CO_2 inversions based on atmospheric data taken from 16 European stations for 2007. Different data-driven error structures in the prior error covariance were assessed, and optimized biospheric fluxes were retrieved and post-processed at various temporal and spatial scales for further evaluation. In this part we discuss the fitting performance of the inversion system, and we detail the comparisons between our flux estimates at grid, national and continental scales against eddy covariance data and reported flux estimates from previous studies. Finally we discuss how sensitive flux retrievals are in the presence of erroneous representation of the fossil fuel fluxes, and the site selection.

5.4.1 Goodness of fit

Site-specific misfits show a reasonable fit to the atmospheric data. Nevertheless in 4 cases (CBW, HEI, JFJ, and KAS) site-specific χ_c^2 values were found to be larger than 1 (see also Table 5.3), indicating that either the model-data mismatch errors were chosen too small, or the spatiotemporal resolution of the flux model is too coarse compared to the biosphere fluxes and therefore small scale variations are not resolved (Rödenbeck et al., 2003). In fact this seems to be the case for the JFJ and KAS sites as those are high altitude sites with steep cliffs. In such a complex terrain the atmospheric circulation is hard to be simulated from the transport models. Regarding CBW and in particular HEI, those are polluted sites and it would be reasonable to assume larger model-data mismatch uncertainty since the model is too coarse to resolve the fossil fuel emission patterns. One could argue that using higher spatial resolution to couple fossil fuel fluxes with transport models might reduce the model-data mismatch uncertainties, and hence improve posterior fluxes. To investigate that, we performed a forward run at coarser (0.25º) and higher (1/12º lat. X 1/8º lon.) spatial resolution using only the fossil fuel emissions. As we use a Lagrangian transport model, fluxes at higher resolution than that of the meteorological fields can be used such that the simulated fossil fuel signals contain more spatially detailed information (Lin et al., 2003). The derived concentration signal was subtracted from the observations and subsequently an atmospheric inversion was performed. We report no significant differences between the retrieved fluxes indicating that simply increasing the spatial resolution to about 10 km is not enough to correctly represent the fossil fuel distribution.

The reduced χ_r^2 values in our study (between 1.08 and 1.25) are larger than those found by Tolk et al. (2011) where values between 0.34 and 0.78 were found for their pixel based inversion, indicating a more conservative choice for their model-data mismatch errors. Even lower values were reported in the study by Peylin et al. (2005) with values ranging from 0.01 up to 0.6 depending on the assumed correlations. χ^2 values from Zhang et al. (2015) were within a range of 1 to 4, but were modified by inflating the error covariances through an iterative procedure, resulting in χ_r^2 values comparable to ours. Concluding, the χ_r^2 values give confidence that the assumed prior uncertainties are well defined. We note though that examining the χ_r^2 values is not always a sufficient metric to evaluate the defined uncertainties (Michalak et al., 2005; Chevallier, 2007).

5.4.2 Validation against eddy flux measurements

At the local scale the inversion shows ability to capture the observed flux variability at monthly scale, as shown for the B2 case (Figure 5.6). The residuals between posterior model and eddy covariance flux-data for monthly and site averaged fluxes show a range of misfits not exceeding 1.04 gCm^2day^{-1} which is comparable with Broquet et al. (2013), where misfits up to 1.5 gCm^2day^{-1} were found using 6 years of data (2002-2007). Of note is that the estimated carbon uptake agrees well with the estimated uptake for 2007 in Broquet et al. (2013) (within the uncertainty range). However, in contrast to the synthetic inversion, the real data inversion showed a larger monthly averaged posterior bias equal to 0.40 gCm^2day^{-1} compared to the -0.04 gCm^2day^{-1} for the synthetic case in chapter 4. The poorer performance in terms of bias compared to the synthetic case is presumably mainly caused by the representation error. In the synthetic inversion we created a true flux field at the same spatial resolution as the posterior flux estimates, and sampled this true flux distribution at the specific eddy covariance measurement location. This does not include any spatial representation error of the EC measurements (footprint about 1 km) with respect to the spatial resolution of 25 km at which the fluxes are used within the inversion. A further cause for this poorer performance is related to the transport error, as in the synthetic case the same transport was used to create the synthetic observations and to perform the inversion, while in the real data inversions the observed atmospheric mole fraction are a result of real transport which can only be approximated with the transport model used for the inversion.

Differences between posterior flux retrievals and observed NEE fluxes at the eddy covariance stations are clearly driven from the crop sites. The good agreement between posterior inverse flux estimates and fluxes measured with the eddy covariance technique at non-crop sites can be attributed to the relatively stable, within the year, land condition. Contrastingly, crop areas are subject to human activities throughout the year. Soil enrichment with organic fertilizers, irrigation and harvesting, can severely influence the carbon balance of the local ecosystem. Thus the poor performance between inverse estimates and eddy covariance flux measurements at crop sites can be linked to the extensive anthropogenic influence on those areas. Further it is worth to mention that atmospheric concentrations implicitly contain more components than just the NEE signal e.g. fire emissions. Such emissions are captured in the atmospheric observations (representative scale of hundreds of km) but might not be captured from the eddy covariance flux measurements which they have a very short representative scale of around 1 km.

5.4.3 Reliability of European flux estimates

5.4.3.1 Mismatch in bottom-up and top-down methods

Table 5.4 Results from Jackknife delete-1 statistics for VPRM estimated domain-wide NEE for different vegetation classes and for all of the land area. The uncertainty in NEE from all land area was derived assuming independence in the vegetation class specific uncertainties. Note the strong asymmetry between the fraction of land area covered by the different vegetation classes and the number of eddy covariance sites used, indicating over/under representation: for example 8 crop sites represent 51% of the land area, while 15 grassland sites represent 5.6% of the land area of Europe.

	NEE [GtC/y]	NEE uncertainty [GtC/y]	Number of sites	Fraction of land area [%]
Evergreen forest	-0.165	0.039	16	16.5
Deciduous forest	-0.174	0.020	5	4.4
Mixed forest	-0.025	0.176	2	8.4
Open shrub[a]	-0.201	-	1	13.8
Savanna[a]	-0.012	-	0	0.3
Crop	-0.443	0.502	8	51.0
Grass	0.059	0.026	15	5.6
Total	0.960	0.536	47	100

[a]Uncertainties for open shrubland and savanna could not be derived due to the lack of representative eddy covariance sites

Of note is the strong flux correction when using a-priori fluxes from VPRM with an uptake of 0.96 GtC y^{-1} compared to the 0.3 GtC y^{-1} after the inversion. The large correction of about 0.66 GtC y^{-1} corresponds to roughly twice the prior uncertainty. We note that VPRM is a diagnostic model which uses simple light use efficiency and respiration equations and MODIS indices, with parameters optimized to match hourly observations of NEE fluxes (Mahadevan et al., 2008). It does not account for land management and land use changes (i.e. crop harvest, deforestation), thus it will estimate a strong sink even for lands that have been harvested, with the respiration fluxes resulting from the use of the harvest (e.g. as food) not included. Those so-called lateral carbon fluxes, that are seen by the atmospheric inversion, account for approximately 0.165 GtC y^{-1}) of the prior-posterior flux difference (Ciais et al., 2008). The rest of the difference of about 0.5 GtC y^{-1} might be related to local characteristics of eddy covariance sites, which VPRM is not able to represent. Spatial variations of NEE from VPRM are driven by those of EVI, which is used at a spatial resolution of 1 km. For example, a crop field with typical dimensions of 100 m – 200 m surrounded by other fields with different crop rotation (and differing phenology) are hard

to represent with 1 km resolution EVI (even with the highest possible resolution of 250 m for MODIS reflectances). To quantitatively assess the impact of this representation error in combination with the selection of sites used for the VPRM optimization, the annual domain wide C-budget from VPRM was recalculated after omitting one site per vegetation type at a time and optimizing the VPRM parameters (Jackknife delete-1 method). Detailed results are shown in Table 5.4. The derived Jackknife standard error amounted to 0.54 GtC y^{-1}, with a dominant contribution from the cropland vegetation class (0.50 GtC y^{-1}). This uncertainty can fully explain the mismatch between the a priori and the posterior fluxes, and it emphasizes the importance of site selection and site representativeness in up-scaling local eddy covariance measurements to larger regions.

The estimated uncertainty for VPRM fluxes based on jackknifing is larger than the prior uncertainties assumed for the atmospheric inversions. Hence, one could argue that the prior fluxes using VPRM (which indicate a too strong sink) combined with a too small prior uncertainty in the inversion leads to erroneous posterior flux estimates. However the optimized biogenic fluxes from all inversions converge at the annual and domain-integrated scale. A particular example is that of the B2 inversion. Even though the GBIOME-BGCv1 fluxes differ greatly from those produced by VPRM, this inversion is fully in line with the results from the rest of the inversions, indicating that the optimized flux estimates are not biased by the a priori flux fields but instead are driven by the atmospheric data.

5.4.3.2 Sensitivity to anthropogenic emissions

Another source of biospheric flux misrepresentation is the fossil fuel inventories. As mentioned in section 5.2.3 we do not allow for corrections in anthropogenic emissions, as they are assumed to be better known than the terrestrial fluxes. An overestimation/underestimation in anthropogenic emissions will thus lead to a stronger/weaker biospheric sink in atmospheric inversions. The anthropogenic emissions we use are 0.32 $GtCy^{-1}$ (27%) lower for the EU-12 countries compared to those used by Rivier et al. (2010) (1.2 $GtCy^{-1}$). Peylin et al. (2011) estimates the difference between national totals for the different emission inventories to be around 10%. In a study by Ciais et al. (2009) uncertainties of total fossil-fuel CO_2 emissions in the European Union 25 member states were estimated to 19%, based on four different emission inventories. For the EU-25 countries, EDGAR emissions were found to be 12% larger than the mean of the GAINS (Greenhouse Gas and Air Pollution Interactions and Synergies), UNFCC (United Nations Framework Convention on Climate Change) and CDIAC (Carbon Dioxide Information Analysis Center) inventories (Ciais et al. 2009, table 2). Sensitivity tests with increased prior fossil fuel emissions showed that the added fossil fuel increases the estimated uptake by almost 50% relative to the added anthropogenic emissions. Taking an extreme scenario where the fossil fuel emissions are increased by 17% or 0.3 GtC y^{-1} (resulting in 1.77 GtC y^{-1} compared to 1.47 GtC y^{-1} total emissions for EU-domain), we estimate a European carbon sink for the B1 set up of -0.51 ± 0.17 $GtCy^{-1}$ compared to -0.38 ± 0.17 GtC y^{-1} for the standard B1 case. Thus the additional assumed fossil fuel emissions increased the estimated uptake by 0.13 $GtCy^{-1}$, which is about 44% of the added anthropogenic emissions. The fact that the resulting increase in the biospheric sink does not fully correspond to the increase in assumed emissions is likely a result of the sparse network, where emissions from regions further away from the measurement sites are not fully registered in the simulated mole fractions.

5.4.3.3 Sensitivity to site selection

Uncertainties in vertical mixing and especially in the nocturnal boundary layer (Gerbig et al., 2008) should be carefully addressed as they might lead to erroneous estimations of the carbon uptake. Typically in atmospheric inversions the model-data mismatch error (measurement error covariance) accounts also for uncertainties due to the transport (i.e. wrong representation of the nocturnal boundary layer). The set of network stations includes 6 mountain stations, for which we use night-time observations (day-time for non mountain stations) as these measurements are considered to be representative for the free troposphere. Errors can be introduced if the measurement height assumed in the transport model is within the modeled nocturnal stable boundary layer while in the real world it is not, which would lead to an overestimation in the simulated CO_2 signals from respiration or vice versa. In the inversion this would be compensated by introducing stronger uptake fluxes to match the observed CO_2 time series. In order to investigate whether our results are influenced by the use of mountain stations, we performed an additional inversion using the B1 error structure, but excluding all these stations. The resulting sink in Europe was found to be -0.41 ± 0.17 $GtCy^{-1}$ which is fully in line with B1 inversion using all sites, suggesting that our estimates are not biased due to misrepresentation of the mountain stations.

5.4.3.4 Retrieved fluxes and comparison to previous inverse estimates

The retrieved spatially resolved fluxes showed a sensitivity in their spatial patterns to the a priori error structure, specifically to the inclusion of a bias component, as indicated by differences between the B1 and S1 cases. Such differences were not identified in the synthetic experiment in chapter 4, however there, a much larger spatial correlation length scale was assumed. In the synthetic inversions the long correlation length (766 km at the zonal and 411 km at the meridional direction) drastically reduces the effective number of degrees of freedom, forcing the fluxes to be smoothly corrected, regardless of the use of the bias component. In the real data inversions the shorter correlation length (around 100 km), combined with the required larger error inflation (compared to the synthetic inversions) for the B1 and B2 cases, increases the effective number of degrees of freedom. By using a bias component (S1, S1a, S1b cases) or by using the hyperbolic correlation shape (S3) with stronger large-scale correlation, instead of inflating the spatiotemporal error component, fluxes remain less flexible at gridscale.

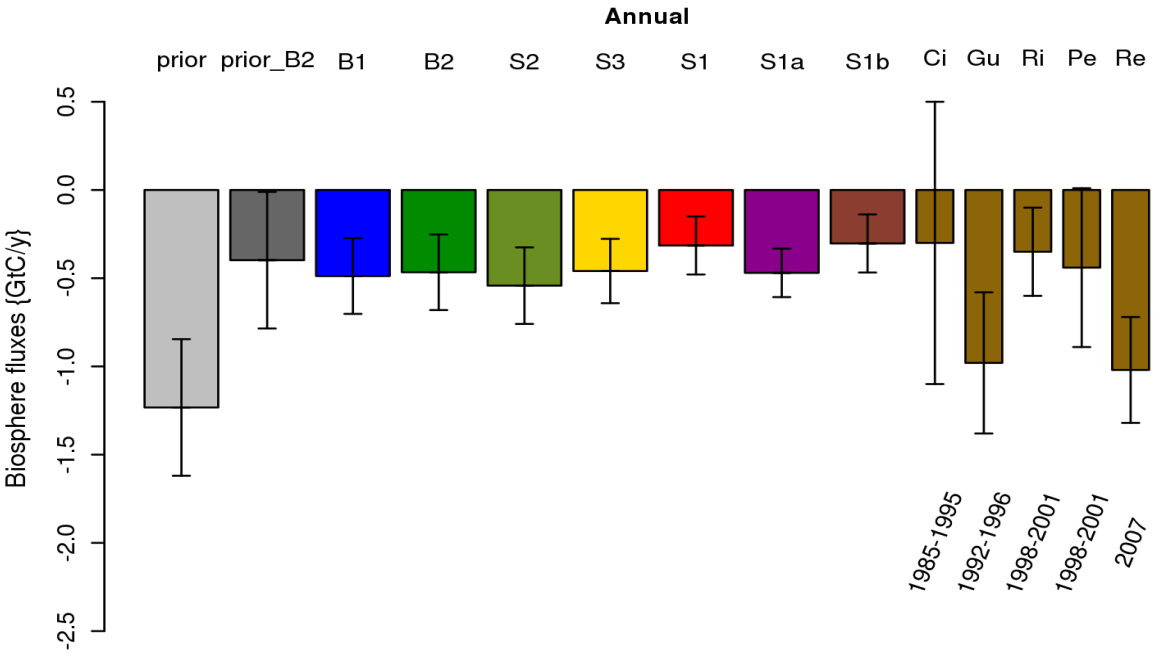

Figure 5.7 Annual European biogenic CO_2 fluxes in $GtCy^{-1}$ for the different inversions and comparison to previous studies. Fluxes are upscaled to the TransCom EU domain. Labels of the references are as follows: Ci : Ciais et al. (2000); Gu : Gurney et al. (2004); Ri : Rivier et al. (2010); Pe : Peylin et al. (2013); Re : Reuter et al. (2014). Periods for the inverted fluxes are given below the flux estimates.

Our knowledge regarding annual CO_2 flux estimates for Europe is still highly uncertain, in part due to the limited number of regional inversions focusing on this domain. Flux estimates from previous studies, mainly global inversions, show a wide range (Figure 5.7). We estimated an annual European carbon sink (ranging between -0.23 ± 0.13 and -0.38 ± 0.17 GtC y^{-1} for the different inversion scenarios, Figure 5.4 d)), which is however representative for a smaller European region compared to the TransCom European region typically used in other studies. The up-scaled flux estimates (see also section 5.2.4) for the TransCom EU region have a range of -0.30 to -0.49 GtC y^{-1}. Ciais et al. (2000) estimated a European sink of -0.3 ± 0.8 GtC y^{-1} for the target period 1985-1995, however in contrast to our study they used a global system and a gap filling algorithm since 42% of the observational data were missing. A recent study from Peylin et al. (2013) computed the mean European sink for the period 1998-2001 to be -0.44 ± 0.45 GtC y^{-1} by utilizing eleven different global inversion systems. Gurney et al. (2004) performed also global inversions and found the mean European annual fluxes for 1992 – 1996 period to be -0.98 ± 0.4 GtC y^{-1} which is larger compared to our estimations. Moreover, our results for the mean net monthly fluxes over Europe agreed very well with Rivier et al. (2010) who estimated for the 1998-2001 time frame using five different transport models in their inversion that the maximum seasonal uptake occurs in July and lies between -10 and -80 $gCm^{-2}month^{-1}$, while our results show maximum uptake in June with a range of -33 to -37 $gCm^{-2}month^{-1}$ for the different inversion cases. We note that the annual flux differences between our flux estimates and those from other studies may be also caused due to the interannual flux variability. Nevertheless this should not be expected to critically drive those differences since posterior uncertainties found to be larger than interannual variations (Broquet et al., 2013) making the significance of the variations questionable.

A recent study from Reuter et al. (2014) based on inversions using satellite observations estimated the carbon budget for the TransCom European region. For the year 2007 the sink was found to be -1.1 ± 0.30 GtC y^{-1}, much larger compared to most of other inversion estimates using ground observations. However Feng et al. (2016) linked the increased uptake when using satellite measurements to potential observation biases and to the emission spatial patterns. Further Feng et al. (2016) highlighted that the large European uptake is related up to 60-90 % from systematically higher modeled CO_2 fluxes transported into Europe from regions outside of the domain. As this looks to be a problem related with column measurements this is not the case in our study since ground observations were used. In addition we use the two step inversion scheme which limits the influence from the far field as we calculate the concentration signal from outside the domain and subtract that from the observations. Whilst the flux uncertainties outside the domain are not propagated, still they can be expressed as uncertainties in the observation space. However if biases introduced from the global inversion to the fluxes outside of the domain, then regional flux estimations may differ.

At national scale we can compare our results to those obtained by Meesters et al. (2012) for the Netherlands, who estimated the annual national carbon sink to about -0.017 ± 0.004 $GtCy^{-1}$. Our estimations are very close, with a range of – 0.012 ± 0.004 $GtCy^{-1}$ (S1 inversion) to – 0.014 ± 0.005 for the B2 inversion. Of note is that the carbon budget estimates for Netherlands agree remarkably well despite the substantial differences between the two studies: Meesters et al. (2012) used an inversion scheme that solves for scaling factors of the gross prior fluxes. Spatial correlations of 100 km were assumed but only for photosynthetic fluxes within the same land use class. In addition the domain of interest (Netherlands) has a stronger constraint as four stations located within the domain were used, while our inversion only uses one station (CBW), with the rest of the stations being at least 360 km away (WES). Both studies assume approximately the same fossil fuel emissions (0.051 GtC y^{-1} vs. 0.053 GtC y^{-1} in Meesters et al. (2012)).

5.5 Conclusions

This study is a follow up work from Kountouris et al. (2016a). In this second part, the inverse modeling framework was deployed using real atmospheric data from 16 stations in Europe, to infer biospheric carbon fluxes. Different prior error structures were assumed to investigate how sensitive posterior fluxes are. The results are validated and compared at different temporal and spatial scales. Satisfactory agreement was found when posterior inverse flux estimates were compared against eddy covariance observations at local scale, as well as against previous studies at national and continental scales, which gives us confidence for our carbon flux estimations. We calculated a sink for the European continent which amounts of -0.23 ± 0.13 GtC y^{-1} to -0.38 ± 0.17 GtC y^{-1} depending on the assumed prior error structure.

A special effort was also made to avoid potential biased flux estimations due to site selection (i.e. heavily polluted sites, or sites that are within the nocturnal boundary layer) by performing inversions using different network configurations. We did not observe any significant impact at least for monthly and annual scales. Further we studied also how sensitive biospheric carbon fluxes are, when wrong fossil fuel emissions are assumed. We found that due to the network sparseness the fossil fuel emissions are not fully captured in the simulated mole fractions which may bias the flux estimates.

Chapter 6

Assessing the ICOS current and future atmospheric network using multiple inversion systems

Abstract

We quantitatively assess the European Integrated Carbon Observing System (ICOS) atmospheric network using two different inversion systems, the TM3-STILT system from Jena and the inversion system from Vrije Universiteit Amsterdam (VUA) which is based on the Ensemble Kalman Filter method. The uncertainty reduction for biosphere – atmosphere exchange fluxes was used as an assessment metric at seasonal and annual scales. To ensure realism in the network assessment, a data driven prior error structure was used with a common protocol for both inversion systems. Differences were found between the two systems regarding the magnitude of the uncertainty reduction as well as the spatial distribution, where TM3-STILT shows a more localized effect. The average domain-wide uncertainty reduction using the full atmospheric network as it is envisioned for the future was found to be 28% and 14% for the TM3-STILT and VUA systems, respectively, whilst significant improvement was realized also at country scale. Furthermore, we assessed the impact of biases in atmospheric measurements on retrieved biosphere-atmosphere exchange fluxes, and found that measurement biases of 0.1 ppm (the World Meteorological Organization (WMO) recommended compatibility goal) are compatible with uncertainty targets for the retrieved spatially resolved fluxes on annual to interannual time scales. However, at country scale measurement biases do not significantly impact the retrieved fluxes unless they exceed values of about 0.4 ppm.

6.1 Introduction

As climate change is one of the most challenging problems that humanity will have to cope with in the coming decades, long-term observations of the greenhouse gas emissions are needed that help to evaluate mitigation activities aiming at reducing greenhouse gas emissions. With that purpose, atmospheric inverse modelling has been widely used to infer carbon dioxide (CO_2) surface fluxes both at global (Tans et al., 1990; Gurney et al., 2003; Mueller et al., 2008; Rödenbeck et al., 2003) and regional scales (Gerbig et al., 2003; Peylin et al., 2005; Broquet et al., 2013). For that, flask and/or continuous in situ data of CO_2 dry mole fractions are used in combination with an atmospheric transport model, running in inverse mode to optimize prior flux estimates.

Despite the great advance of atmospheric modelling results between different studies still show a large spread. For example the uncertainty of European fluxes among 11 different inversion systems is found to be as large as the annual estimated fluxes (Peylin et al., 2013). The use of regional inverse systems with a denser atmospheric network is expected to provide more accurate flux estimates. For that, ground based measurement stations have been developed over the last two decades in the globe monitoring atmospheric CO_2 dry mole fractions. A more

coordinated initiative is the Integrated Carbon Observing System (ICOS), a pan European research infrastructure which aims to provide European – wide greenhouse gas data by establishing a monitoring network. Whilst the sampling network is continuously expanding, the choice of a measurement location impacts the uncertainties of the estimated fluxes. Atmospheric inverse modelling is not only used to derive optimal flux estimates but also to investigate and infer the optimal network which can either maximize the information of specific target quantities (i.e. annual or monthly CO_2 fluxes at national or domain wide scales) or minimize a chosen uncertainty metric (i.e. the posterior flux uncertainty). The great advantage of this framework is the ability to assess also hypothetical networks where no observations exist as the posterior uncertainty does only depend on the prior and model – data uncertainties and not on the observations themselves.

Early studies on network design carried out by Rayner et al. (1996) and Rayner (2004) and later by Ziehn et al. (2014) and Nickless et al. (2015) used a number of optimization methods like simulated annealing, incremental approach and genetic algorithms to infer an optimal network which minimizes a chosen metric. Other regional network design studies used prespecified station combinations to study their impact on the regional flux balance and its spatial distribution (Lauvaux et al., 2012) or on the posterior flux uncertainty (Kadygrov et al., 2015). For network assessments the posterior uncertainty or the uncertainty reduction, defined as the ratio between the difference of prior and posterior over the prior uncertainties, of different candidate networks can be used to assess their potential. As the posterior uncertainty depends on the measurement uncertainty, uncertainty of the modeled transport, and on the a priori uncertainty of the spatiotemporal flux distribution, all those sources of uncertainty need to be carefully quantified. Especially the latter has large impact on the outcome of network design (Nickless et al., 2015).

Describing the prior error structure in a direct and quantitative way is challenging, and leaves room for assumptions and ad hoc solutions. Some studies assumed that prior flux uncertainties are independent for different spatial regions or grid elements (Patra 2002; Rayner 2004; Ziehn et al., 2014) or assumed flux error correlations derived from sensitivity tests (Lauvaux et al., 2012; Nickless et al., 2015). The choice of the spatial error correlation length is of high importance as it controls how far the information is spread. Large spatial correlation scales serve two purposes: 1) they regularize the inverse problem as they reduce effective number of degrees of freedom and (with unchanged diagonal elements) increase the spatially integrated prior uncertainties, and 2) they artificially increase the area that the network is sensitive to. This would produce a smoother posterior spatial flux distribution and an uncertainty reduction which expands far away from the station however the question arises whether this is realistic.

Another issue rises with the expansion of the network: often different institutions are responsible for setup and operation of the stations, which can lead to different setups and potential calibration differences between different stations. Although random errors will average out at least for longer time scales, systematic errors can occure between measurements of different laboratories (Masarie et al., 2001). The WMO recommendations for GHG measurement require compatibility between different stations of 0.1 ppm for CO_2 and 2 ppb for CH_4 (GAW report No.213, 2013). With the increasing density of stations over the continent, where the temporal and spatial variability is much larger for higher resolution modeling frameworks, the corresponding representation errors in inverse transport models become an issue. These compatibility goals put forward by WMO need to be motivated by the targeted uncertainty in flux estimates, and ultimately by the underlying scientific questions. The impact of undetected

biased observations has been examined in a previous study focused on inversions using satellite observations (Patra et al., 2003). They found that when surface observations are combined with the satellite data in the inversion then posterior flux estimates significantly differ. For measurement bias of about 1 ppm, posterior fluxes aggregated to TransCom-3 regions are found to differ from the "unbiased" inversion up to 2 $GtCy^{-1}$. Rödenbeck et al. (2006) also performed a study on systematic errors in ground atmospheric measurements and how they affect retrieved flux estimates from inversions. They estimated the bias between flask pair measurements and coincident hourly mean values from a continuous analyzer at Samoa station. This bias was then introduced to a limited number of sites and thereinafter a global inversion at a coarse resolution of 4^o lat x 5^o lon was realized. They confirmed almost the same difference as in Patra et al. (2003) on the posterior flux retrievals. Systematic differences on the order of 0.2 ppm lead to regional flux differences on the order of 0.1 $GtCy^{-1}$. Despite of the small impact of a measurement bias which is twice the WMO recommendations, one should be aware that higher resolution inverse models could show different impacts. As the resolution increases and the uncertainties in transport models become smaller, measurement errors will become more important.

This study focuses on high resolution regional CO_2 inversions of observations from the ICOS atmospheric network to retrieve biosphere-atmosphere exchange fluxes, as current inversion systems are not yet capable for estimation of fossil fuel CO_2 emissions. The aim of the study is to provide insight into the uncertainty reduction using two mesoscale inversion systems with different transport models, but similar a priori flux error structure, similar selection of observations, and similar assumptions on transport model and measurement errors. The inversion systems differ in their transport representation, prior flux, and optimization technique. Further we assess the impact of bias errors in greenhousegas observations made at different stations on retrieved fluxes using one of the regional inversion systems. The scientific question we adress is if measurements that do not comply with the WMO reccomended compatibility are expected to be useful as input data for regional inversions over continental areas.

The chapter is structured as follows. Section 6.2 is dedicated to the methods and gives an overview of the inversion systems the setup and the joint protocol used to harmonize the simulations made by the different inversion systems. The method followed to assess the bias errors at station measurements is also reported in this section. In section 6.3 we present our results and discussion follows in section 6.4.

6.2 Methods

6.2.1 Atmospheric inversions and uncertainty reduction as a tool for network design

Atmospheric inversions use observed time series of atmospheric dry mole fractions from a network of stations, in combination with a transport model to infer optimal fluxes, i.e. fluxes that result in simulated dry mole fractions that are optimally close to the observations. They rely on the Bayesian update of a prior guess for the fluxes. This can be expressed in a cost function J that needs to be minimized:

$$J = (Kx - y)^T Q_y^{-1} (Kx - y) + (x - x_p)^T Q_x^{-1} (x - x_p)$$

6.1

The vector y contains observed dry mole fractions at different times and locations, K is the transport model (as the transport is linear for conserved tracers, K is a matrix), x is the vector of fluxes (e.g. can be the fluxes themselves or scaling factors) at all different times (e.g. 3-hourly) and locations (e.g. every 0.5 degree lat/lon), χ_p is the a priori guess for the fluxes, and Q_y and Q_x are the model-data mismatch uncertainty (due to the discrete representation of the fluxes, to the transport model and to the measurement errors) and the prior uncertainty, respectively. Assuming that these uncertainties have Gaussian distributions, they can be expressed in a form of covariance matrices. Note that the vector χ is often called "control vector" or "state space", as it contains all the adjustable elements which control the biosphere-atmosphere exchange fluxes. The cost function thus consists of one term describing the model-data mismatch and one term describing the mismatch between what we think as "true" flux and the prior flux. Minimizing Eq. 1 leads to the posterior estimate for the fluxes F:

$$F(y) = (K^T Q_y^{-1} K + Q_x^{-1})^{-1} (K^T Q_y^{-1} y + Q_x^{-1} F_{prior})$$

6.2

The posterior uncertainties can be derived from Eq. 3 :

$$Q_{x,post} = (K^T Q_y^{-1} K + Q_x^{-1})^{-1}$$

6.3

The posterior flux uncertainty depends on the transport, which decides how strongly the flux elements have been "seen" by the observation locations (*K* contains the sensitivity of dry mole fraction observations to upstream fluxes). The posterior uncertainty also depends on the uncertainties for model-data mismatch and prior fluxes, but it does not depend on the prior fluxes, nor on the observations.

Targeted quantities can be assigned from the estimated surface flux fields by defining an aggregation scalar operator *W* following also Rodgers (2000):

$$F_{post} = W x_{post}$$

6.4

As we are interested in national annual totals the *W* operator corresponds to annual mean flux, averaged over a given national domain. Similarly, the full error covariance matrix $Q_{\chi,post}$ of posterior fluxes can be aggregated to the standard deviation of the posterior uncertainty $\sigma_{F,post}$ of the target quantity:

$$\sigma_{F,post} = \sqrt{W Q_{x,post} W^T}$$

6.5

For network assessments, the posterior uncertainty or the uncertainty reduction of different candidate networks can be used to assess their potential. Note that for areas not "seen" by the network, i.e. where footprint is zero, there is formally no potential for atmospheric measurements to constrain fluxes. An exception occurs if appropriate spatial or temporal correlations exist in

the a priori flux uncertainty. Such areas appear to be indirectly constrained by directly constrained (and error-correlated) fluxes elsewhere. In other words, fluxes in areas with gaps in the footprints can be constrained as long as the spatiotemporal correlation length is sufficiently large. Thus specifying the spatiotemporal correlation of errors is crucial, as it spreads the information on top of what is given by atmospheric transport alone. On the other hand, the number of unknowns (wall-to-wall fluxes with a high spatial and temporal resolution) easily exceeds the number of observations (hourly time series at each network site), which means that the problem is underdetermined. Thus the smoothing and extrapolation provided by spatiotemporal correlations in the a priori error helps constraining fluxes beyond the near vicinity (footprint) of stations.

6.2.2 Inversion Systems

6.2.2.1 Jena Inversion System TM3-STILT

The Jena inversion system with the two step inversion approach was deployed (Rödenbeck et al., 2009). The inversion system uses the global TM3 atmospheric model (Heimann and Körner, 2003) and the STILT model (Lin et al., 2003) for regional simulations. For more details the reader is referred to Rödenbeck (2005) and Kountouris et al. (2016a).

For the current study only the regional model was used driven by ECMWF short term forecast fields (Trusilova et al., 2010) . The control or state space (variables to be optimized) corresponds to NEE fluxes, such that the system solves for additive corrections to three hourly a priori NEE fluxes. The optimization scheme is an iterative matrix inversion, comparable to typical 4D-variational schemes used in weather prediction. To derive posterior uncertainties, the inversion system uses a Monte Carlo method, where 40 ensembles of realizations of prior error and model-data mismatch errors are generated, and the inversion is repeated for each ensemble member, resulting in posterior fluxes that exhibit a spread corresponding to the posterior uncertainty. For each ensemble inversion 70 iterations were realized ensuring the system convergence.

6.2.2.2 Inversion system VUA

The VUA inversion system uses RAMS (the Regional Atmospheric Modelling System) as transport model. This weather prediction model with capability for tracer transport is run in forward mode for an ensemble of flux realizations (created using realizations of prior error), to create an ensemble of synthetic observations.

Inversions are performed only for periods of three months: MAM, JJA and SON 2007, as winter fluxes are very small and hard to correct by the inversion (Meesters et al., 2012). The control space consists of scaling factors for GPP and respiration for each cell-month combination, and one bias factor (following a respiration flux shape) for the whole season and domain. For the control vector, an ensemble of 100 random members is set up, whose statistics satisfy the protocol (see section 6.2.3). Forward runs with RAMS yield “observations” for each member. Inversion is done with an ensemble Kalman filter (EKF), which processes the central concentrations observation-by-observation, thus transforming step-by-step the prior to a posterior flux ensemble (Peters et al., 2005).

Since the ensemble has the imposed statistical properties only in approximation, relations between the variations in the fluxes and concentrations may occur not only by causality, but also by coincidence, causing spurious "improvements" of the fluxes when the EKF is applied. To prevent those improvements the following measures were taken:

(1) Earlier observations are not allowed to improve later fluxes.
(2) Observations are not allowed to improve fluxes more than 1000 km away ("localization").
(3) "Dynamic localization" was used where an adjustment of an observation is only accepted if it reduces the scaling-factor's standard deviation by at least 5 % (Zupanski et al., 2007).

Unfortunately, the second and third measure creates a spurious dependence of the results on the order in which the observations (stations) are processed. This dependence is quite weak for small station numbers, but it becomes strong with many stations, with late stations having a smaller impact than early ones. To overcome this difficulty, five "reference locations" were defined, being the four corners of Europe and the centre of Germany. For each reference location, a complete inversion is done, with the stations ordered with increasing distance to the reference location. After all inversions are finished, for each pixel, the best of the 5 flux error reductions is chosen as final result.

6.2.3 Joint protocol for regional inversions

6.2.3.1 Observational networks

This study focuses on existing and planned atmospheric stations within ICOS (see also Table 6.1). ICOS current (denoted as ICOS1) and future (denoted as ICOS2) stations were taken from the ICOS Stakeholders Handbook 2012 and 2013. All locations were updated using Google-earth. The overview of the network is given in Figure 6.1. Further to examine how network spatial gaps impact the uncertainty reduction at country scale a network using the ICOS2 configuration but excluding all stations in Germany was assessed. We will refer to this network as ICOS2EG hereafter.

A common model domain for comparing the flux retrievals was chosen with the south-west corner at 35°N, 11°W and the north-east corner at 61°N, 35°E. The uncertainty reduction analysis is focused on the full year 2007.

We note one difference between the station networks used in TM3-STILT and VUA systems. Unintentionally the Plateau Rosa station (PRS) was not considered in the TM3-STILT simulations for the ICOS2 network. However we do not expect our results to be biased. We argue that the number of stations is representative for the statistical analysis and further as it is shown in the results section (see 6.3.1.1 section and Figure 6.4) TM3-STILT uncertainty reduction is of local nature. Hence any bias would occur only at the close proximity of the station. Potentially uncertainty reduction estimations may be biased at country level for neighboring countries like Switzerland and eastern part of France. Nevertheless two more stations are present in this location (JFJ and IPR) ensuring that there is satisfactory observation coverage.

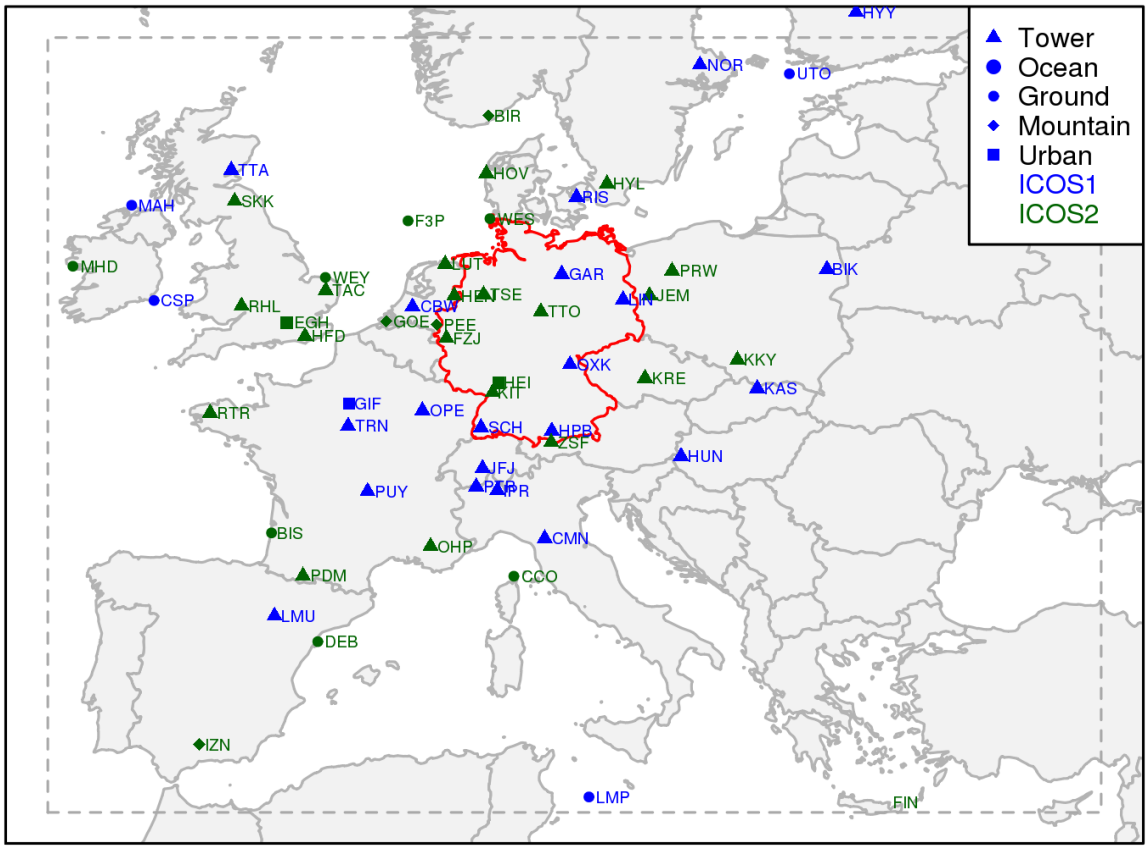

Figure 6.1 Site locations and types for the current atmospheric network (ICOS1, blue symbols) and additional stations included in the future network (ICOS2, green symbols). Borders with the red color indicate the stations (within the borders) excluded in the ICOSEG configuration.

Table 6.1 Current atmospheric measurement sites (ICOS1) and potential locations for the future expansion of ICOS network (ICOS2). Altitude corresponds to the height above sea level, and M_height is the height above ground. Types of stations are coded as follows: tall tower (T), mountain (M), ocean/coastal (OC), urban polluted (UP) and ground (G).

ID	Name	Latitude	Longitude	Altitude	M_height	Type
ICOS1						
BIK	Bialystok	53.23	23.03	183	300	T
CBW	Cabauw	51.97	4.93	20	200	T
CMN	Monte_Cimone	44.19	10.70	2165	10	M
CSP	Carnsore_Point	52.18	-6.36	10	3	OC
GAR	Gartow	53.07	11.44	66	344	T
GIF	Gif_sur_Yvette	48.71	2.15	165	7	UP
HPB	Hohenpeissenberg	47.80	11.01	947	132	T

HUN	Hegyhatsal	46.96	16.65	248	96	T
IPR	Ispra	45.80	8.63	209	65	T
JFJ	Jungfraujoch	46.55	7.99	3580	10	M
KAS	Kasprowy_W.	49.23	19.98	1984	5	M
LIN	Lindenberg	52.21	14.12	93	99	T
LMP	Lampedusa	35.52	12.63	45	10	OC
LMU	LaMuela	41.59	-1.10	570	80	T
MAH	Malin_Head	55.37	-7.34	22	3	OC
NOR	Norunda	60.09	17.48	70	102	T
OPE	OPE	48.48	5.36	392	120	T
OXK	Ochsenkopf	50.03	11.81	1022	163	T
PUY	Puy_de_Dome	45.77	2.96	1465	10	M
RIS	Riso	55.65	12.09	5	125	T
SCH	Schauinsland	47.92	7.92	1205	8	M
TRN	Trainou	47.96	2.11	131	180	T
TTA	Angus	56.56	-2.99	313	222	T
UTO	Uto	59.78	21.38	8	60	OC
ICOS2						
BIR	Birkenes	58.38	8.25	190	40	G
BIS	Biscarros	44.38	-1.23	12	47	OC
CCO	Cap_Corse	42.93	9.35	0	35	OC
DEB	Delta_Ebre	40.73	0.79	0	5	OC
EGH	Egham	51.43	-0.57	60	10	UP
F3P	Noordzee_F3P	54.85	4.73	0	50	OC
FZJ	FZ_Juelich	50.91	6.41	98	120	T
GOE	Goes	51.48	3.78	0	70	G

HEI	Heidelberg	49.42	8.67	116	30	UP
HEN	Hengelo	52.34	6.75	14	70	T
HFD	Heathfield	50.98	0.23	256	135	T
HOV	Hovsore	56.44	8.15	1	116	T
HYL	Hyltemossa	56.10	13.42	115	150	T
IZN	Iznajar	37.28	-4.38	403	5	G
JEM	Jemiolow	52.35	15.28	162	314	T
KIT	Karlsruhe_IT	49.09	8.43	115	200	T
KKY	Katowice_Koszowy	50.19	19.12	0	355	T
KRE	Kresin	49.57	15.08	540	250	T
LUT	Lutjewad	53.40	6.35	1	60	T
MHD	Mace_Head	53.33	-9.90	25	15	OC
OHP	Obs_Hau_Provence	43.93	5.71	640	15	M
PDM	Pic_du_Midi	42.94	0.14	2835	10	M
PEE	Peel	51.37	5.98	10	70	G
PRS	Plateau_Rosa	45.94	7.71	3480	0	T
PRW	Pila_Rusinowo	53.17	16.26	134	320	T
RHL	RidgeHill	48.41	-3.91	204	152	T
RTR	Roc_Tredudon	55.51	-2.84	370	15	T
SKK	Selkirk	52.52	1.14	300	229	T
TAC	Tacolneston	52.37	8.03	70	191	T
TSE	T_Schleptruper_E	51.80	10.53	151	234	T
TTO	Transm_Torfhaus	54.92	8.31	825	279	T
WES	Westerland	52.95	1.12	8	10	OC
WEY	Weybourne	47.42	10.99	15	10	OC
ZSF	Zugspitze	53.23	23.03	2971	10	M

6.2.3.2 Prior error setup

Prior uncertainties for different inversions should be consistent regarding the error structure and the uncertainty for annually and spatially integrated biosphere-atmosphere flux within the domain of interest. The error structure was derived from comparison to flux observations at ecosystem sites, and details it in terms of temporal and spatial correlation and in terms of the local standard deviation (Kountouris et al., 2015).

The TM3-STILT system uses a temporal correlation that corresponds to an exponential decay with a time scale of 31 days. The VUA system uses a time resolution of 1 month, and temporal correlations for lag-times longer than that are represented by a polynomial function. The spatial correlation analysis of model-data differences resulted in correlation length scales of 30-40 km (Kountouris et al., 2015). These scales were derived from a site scale analysis whilst the inversion has a coarser resolution. Chevallier et al. (2012) showed the influence of aggregated distances onto the autocorrelation, hence we assume somewhat larger spatial correlation scale. Therefore spatial correlation scale of arround 100 km was implemented. Diagonal elements of the prior error covariance of 2.27 μmolm-2s-1 were considered, which are consistent with the analysis of flux model – eddy covariance differences, scaled down to account for the difference in spatial resolution of the state vector. Initial assessments of the resulting uncertainty budget for annually and domain-wide integrated fluxes suggest additional degrees of freedom to adjust large-scale and long-term mismatch that is not captured by the analysis presented in Kountouris et al. (2015). In detail the domain integrated annual uncertainty using the abovementioned scales is equal to 0.15 GtCy-1.

The prior uncertainty for annually and domain wide integrated fluxes was harmonized for each inversion system. It is selected to be two times the standard deviation of annual terrestrial flux estimates for 2007 between terrestrial biosphere models taken from the global carbon atlas (http://www.globalcarbonatlas.org). As in Kountouris et al. (2016b) we use a domain-wide and annually aggregated prior uncertainty of 0.3 GtCy-1 which is in line also with the uncertainty assumed by the global TM3 inversion system. Hence the aggregated spatiotemporal error covariance leaves a space for a bias term of 0.26 GtCy-1.

The bias term implementation for the two systems differs. In TM3-STILT system a bias term following a respiration shape was implemented and scaled with a single (annual) factor. The VUA system uses a bias component consisting of seasonal factors that scales respiration fluxes from 5PM flux model (Groenendijk et al., 2011). More specifically seasonal biases assumed to be equal to each other and fully correlated; hence they have a value equal to the ¼ of the annual bias as this was defined in the joint protocol (0.26 GtCy-1).

6.2.3.3 Model-data mismatch error setup

Common model-data mismatch errors as used in Kountouris et al. (2016b), presented also in Table 6.2, were used as representative for current inverse modelling systems and are consistent with those that TM3-STILT uses for the global simulations. The sites are categorized as mountain, ground, tall, near shore or polluted stations and site specific uncertainties were applied accordingly. The uncertainties are effective uncertainties for 1 week observing period considering also an aditional error inflation to account for the fact that model-data mismatch errors are not fully random noise, but they are correlated typically over synoptic time scales. For

example, assuming 14 observations per week (12:00 and 15:00 or 0:00 and 3:00 observational times every day), the inflation factor defined as the square root of the number of observations will be about 3.7.

Table 6.2 Model-data uncertainties in ppm for tall towers (T), ground stations (G), ocean/coastal stations (OC), mountain stations (M) and urban polluted sites (UP). Data weight refers to the factor that model-data uncertainties are reduced to mimic future transport model capabilities.

Data weight	T	G	OC	M	UP
1	1.5	2.5	1.5	1.5	4
0.3	0.45	0.75	0.45	0.45	1.20
0.1	0.15	0.25	0.15	0.15	0.4
0.03	0.045	0.075	0.045	0.045	0.12
Number of sites	28	7	11	8	3

6.2.4 Measurement bias experiment

6.2.4.1 Setup characteristics

To assess the impact of the measurement bias the TM3-STILT system was deployed. First, forward simulations were carried out using BIOME-BGC simulated fluxes (Trusilova and Churkina, 2008) as known truth for CO_2, providing synthetic data at the different observing stations. As we are here primarily interested in biosphere-atmosphere exchange fluxes, fossil fuel fluxes were not included. All simulations (forward and inverse) use hourly measurements with a 6-hours local time window (11:00 to 16:00), except for mountain sites, where night-time data were used (23:00 to 4:00 local time). As this analysis utilizes future ICOS stations (ICOS2), observations are not yet available; we used simulated timeseries as observations with no gaps although this would not be the case in real measurements.

A reference inversion assuming unbiased observations was set up to retrieve posterior fluxes using the synthetic data from 57 stations. Inversions are performed multiple times, each with a bias introduced at a given station. As the resulting 57 flux retrievals (for 57 stations) depend linearly on these measurement biases, subsequent analysis can be used to infer maximum bias errors between stations and between parts of the network (e.g. biases between different countries) that still are compatible with the targeted uncertainties. As target quantities we chose annual domain-wide and at national scale fluxes. Error settings for a priori fluxes and model-data mismatch are chosen according to the joint protocol (see also 6.2.3.2 and 6.2.3.3 sections). We ensure that the impact of measurement bias is thoroughly studied by using a range of values for the model-data mismatch corresponding to current transport models, but also to future model capabilities (with smaller model-data mismatch) as transport error is expected to be reduced for future inversion systems. To achieve this, the TM3-STILT inversion system allows for modification of the relative weight of the data constraint vs. the a priori constraint after the

inversion, by adjusting a parameter in the cost function (Rödenbeck 2005). This parameter was used to reduce the effective model-data mismatch error for the atmospheric stations, by a factor of 0.3, 0.1 and 0.03. The uncertainties in this case, are the product of the uncertainty factor, and the station specific model-data uncertainties (Table 6.2).

6.2.4.2 Targeted uncertainties for flux estimates

Posterior uncertainties in estimated (posterior) fluxes are represented as large uncertainty covariance matrices that describe the uncertainties at the doxel level (doxel = spatial pixel and time step). This includes the covariances of errors in fluxes at different locations and times; given the high spatial (0.25^o) and a daily temporal resolution the covariances have about 50 trillion elements (square of the product of total grid cells and number of time steps). Hence specifing targeted uncertainties for each of those elements would not be feasible. In addition atmospheric inversions should be able to give insights to specific scientific questions and for specific purposes. For example as the countries are obliged to report their national greenhouse gas emmisions, inversion studies should be able to infer uncertainties at country level and annualy integrated. The temporal scale targeted in this study is annual and the spatial scales we focus are on 25 km, national and EU-wide domain.

A typical uncertainty reduction of the Jena inversion system is around 75% for biosphere-atmosphere exchange fluxes aggregated to national and domain-wide aggregation scales; we thus set a targeted posterior uncertainty of 25% relative to prior uncertainties. The bias in the retrieved fluxes due to the biased observations should not exceed this threshold value. At the pixel level we are interested in solving for interannual variations of the biospheric fluxes. We used CO_2 modeled fluxes from 9 terrestrial biosphere models (Sitch et al., (2015)), part of the TRENDY database (http://dgvm.ceh.ac.uk/node/9), and we estimated the interannual variation at pixel level for 30 years starting from the year 1980. The average variation for Europe was found to be 71 gCm-2y-1. To set the target uncertainty we assume a signal to noise ratio equal to S/N = 5. In this case the targeted uncertainty should not exceed a value of 20% of the expected signal (i.e. 14.2 gCm-2y-1).

6.2.4.3 Computation of the flux retrieval bias

The inversion for fluxes F from all observations y (a vector containing all observation times and locations) is given by Eq. 2. Assuming a bias in station i with the modified vector of observations being yi, the retrieved flux will differ from that without bias by

$$\Delta F_i = F(y_i) - F(y) = (K^T Q_y^{-1} K + Q_x^{-1})^{-1} K^T Q_y^{-1} (y_i - y) \qquad \textbf{6.6}$$

These flux differences can be computed by performing one inversion per station, each with a bias added to the pseudo data for that station. The flux bias is linearly dependent on the observation bias $y_i - y$. If we assume statistically independent measurement bias errors for all the different stations, error propagation results in

$$\Delta F = \sqrt{\sum_{i=1}^{N} (\Delta F_i)^2}$$

6.7

for the posterior flux bias. Note that the resulting bias error in posterior fluxes, i.e. ΔF, can be computed for various time scales and either for each spatial grid element or for domain-wide aggregated fluxes.

6.3 Results

6.3.1 Assessment of current ICOS network and comparison of the uncertainty reduction from different inversion systems

In this section we assess the performance of the ICOS1 network (25 stations), using the uncertainty reduction as metric. Results from the two different inversion systems, the TM3-STILT and the VUA system, are compared with the aim to better understand potential similarities or differences.

6.3.1.1 Uncertainty reduction at grid scale

We analyze the uncertainty reduction for different seasons except for winter (as the months DJF did not form a contiguous time series and also due to the small flux values is hard to correct by the inversion), and for the full year 2007. Using the currently existing ICOS1 stations, the uncertainty reduction for seasonal fluxes typically reaches values of around 40% and 50% in the vicinity of observing stations (Figure 6.2). The average annual uncertainty reduction for land pixels only is 12.5% for TM3-STILT, while for the VUA system it is slightly smaller and equals to 10.2% (see also Table 6.3). Although both systems have comparable overall uncertainty reduction, TM3-STILT seems to drop stronger with distance to the stations.

Table 6.3 Seasonal and annual domain-wide averaged uncertainty reduction for both inversion systems and networks.

	Period	MPI (%)	VUA (%)
ICOS1	MAM	10.9	8.8
	JJA	10.2	8.9
	SON	12.3	13
	annual	12.5	10.2
ICOS2	MAM	15.6	10.1
	JJA	14.2	10.3
	SON	16.7	15.1
	annual	17.4	11.8

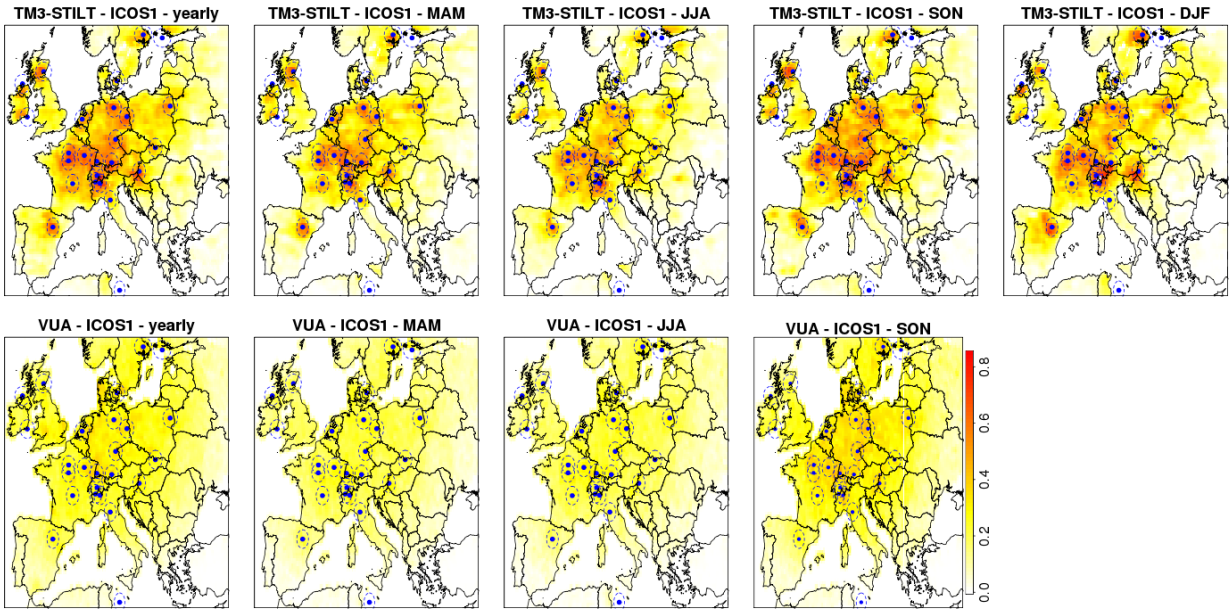

Figure 6.2 Uncertainty reduction maps for ICOS1 network, for different seasons (MAM: March, April, May; JJA: June, July, August; SON: September, October, November) and for the full year 2007 for TM3-STILT (top panels) and the VUA inversion system (bottom panels). Dashed circles around each station indicate the 100km correlation length assumed in the prior error covariance.

Seasonal differences are minor compared to the spatial differences and also smaller compared to the differences between the different station networks (see also Table 6.2 and later section 6.3.1.3). An exception seems to be for fall (SON) where both systems show on average the largest uncertainty reduction 12.3% and 13% for TM3-STILT and VUA respectively. The same result emerges also for the future (ICOS2) network (Figure 6.3). However the increase of the uncertainty reduction from summer to fall is much stronger for VUA system. For both station sets VUA shows a relative increase of around 32%, two times larger than that found for TM3-

STILT (15%). The minimum uncertainty reduction occurs also at different season for the two systems. TM3-STILT for both sets shows the minimum to be in summer whilst VUA in spring.

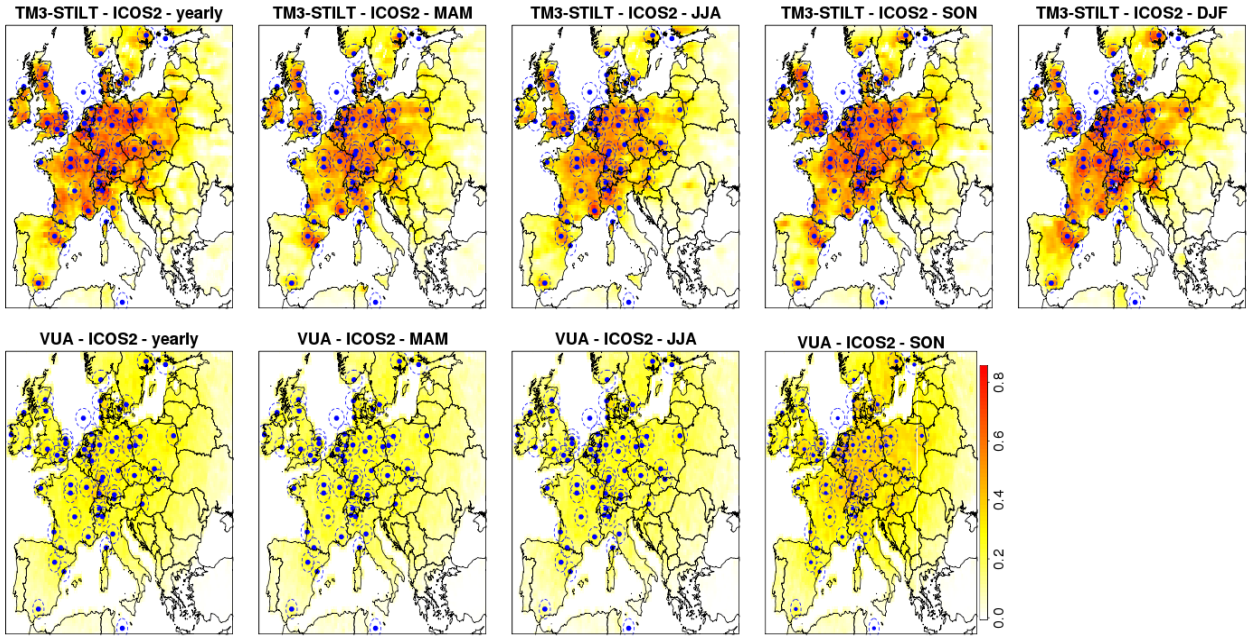

Figure 6.3 Same as Figure 6.2 but for the ICOS2 network.

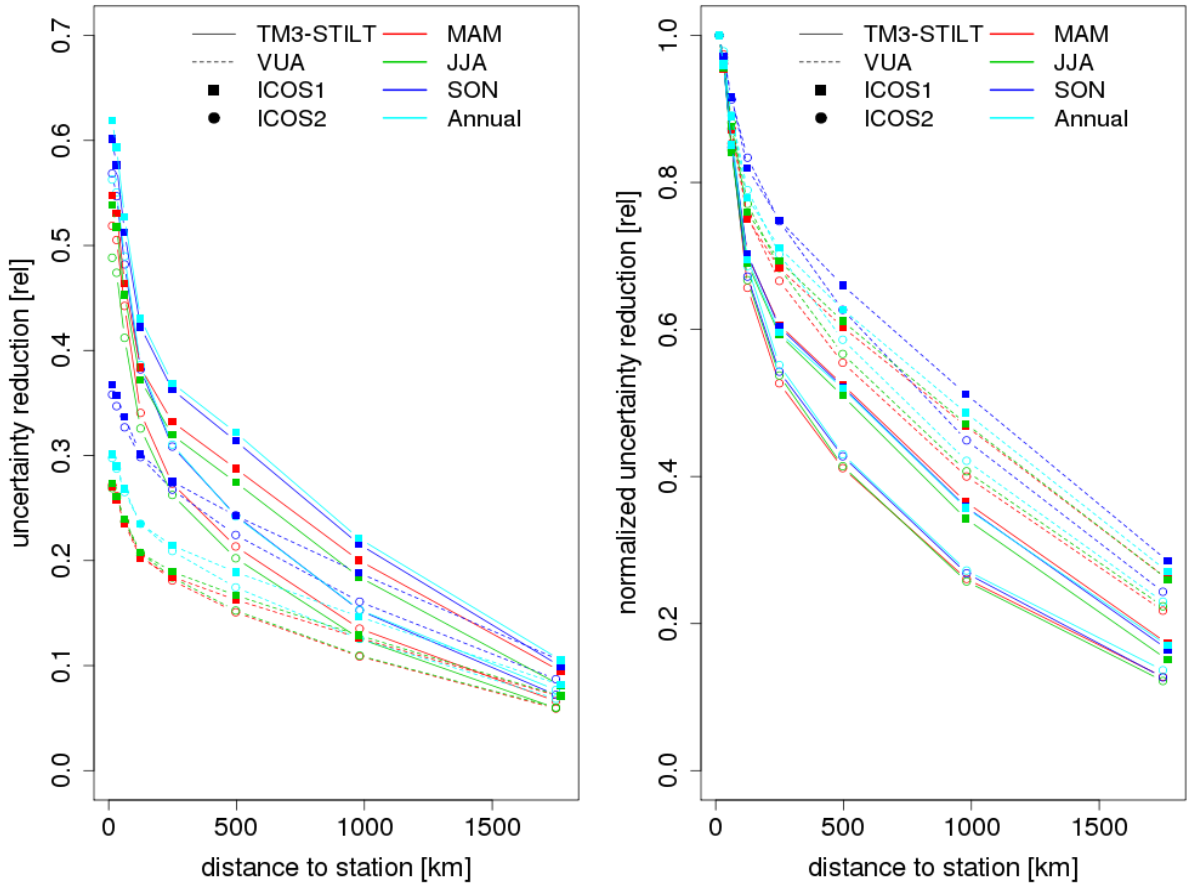

Figure 6.4 Dependence of spatial pattern of uncertainty reduction on distance to the closest station for the different inversion systems, seasons, and network configurations (left panel). The same quantity when normalizing to the respective uncertainty reduction at zero distance is presented at the right panel.

Figure 6.4 shows the average uncertainty reduction for seasonal fluxes at pixel-level as a function of distance to the closest atmospheric station. This analysis reveals the different behavior of the inversion systems. A clear offset is visible between results from TM3-STILT and the VUA system, with overall lower uncertainty reduction from the VUA system. The relative decrease in uncertainty reduction with distance from the stations (normalized by the uncertainty reduction of the corresponding station location) (Figure 6.4, right panel) is stronger for TM3-STILT. This confirms the more localized effect in TM3-STILT which was already identified from the spatial distribution as presented in Figure 6.2. In terms of absolute numbers, TM3-STILT shows an average decrease in uncertainty reduction (seasonally averaged) with distance of around 32% in the first 120 km. Despite same correlations were also assumed for the VUA inversion system, the corresponding decrease is smaller at around 26%, indicating potential differences in the assumed transport between the two systems.

6.3.1.2 Uncertainty reduction at the national scale

Uncertainty reduction at country scale was estimated for those EU28 countries that are represented within the spatial domain as defined in Figure 6.1. Seasonal and annual uncertainty reduction for ICOS1 (the current network) and for both inversion systems is presented in Figure 6.5. The two regional inversion systems show similar patterns in uncertainty reduction for the different countries, with somewhat greater similarity during the fall season ("SON") with the explained variance (squared Pearson correlation) being 0.89, and slightly less similarity during summer ("JJA") with an explained variance of 0.78. Differences in spatial patterns, both, at grid scale (Figure 6.2) or at national scale (Figure 6.5), between different seasons are hard to identify, which suggests that differences in transport patterns (i.e. wind speed and wind direction) are rather small at seasonal scales.

All panels in Figure 6.5 suggest that countries located in central Europe show larger uncertainty reduction values compared to countries located at the eastern and southeastern part of Europe. For example Germany and France have almost the same annual uncertainty reduction (the largest among the European countries) according to both inversion systems. Values were found to be around 75% and 62% for TM3-STILT and VUA respectively. Contrastingly southeastern Europe does not exceed an uncertainty reduction of 10% - 19% (i.e. Bulgaria, Greece).

The countries in Figure 6.5 are ordered by their size. Because of the rather random uncertainty reduction pattern, it is clear that the driver is the country's location relative to the distribution of stations (most of which are located in central Europe), and not the country area. Nonetheless, the number of stations within a country does not necessarily result in a correspondingly large reduction in uncertainty. For example Italy with a network of 5 stations has an uncertainty reduction of 44% and 37% for TM3-STILT and VUA respectively, while Ireland with only 2 stations has 62% and 40% reduction respectively. Furthermore, countries with only 1 station (e.g. Spain, United Kingdom) or even without stations (e.g. Austria, Czech Republic) do have similar or even greater uncertainty reduction compared to countries with a denser network.

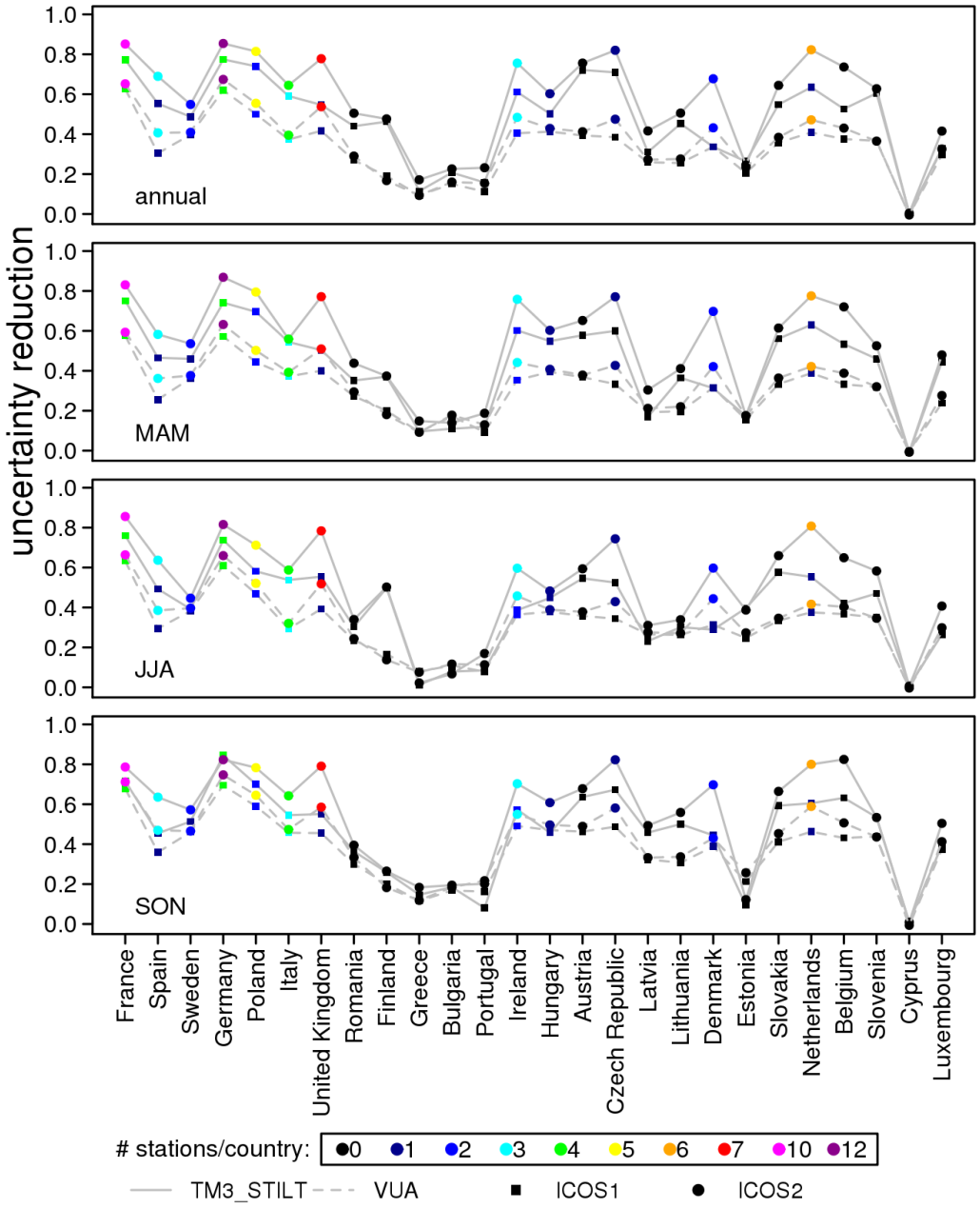

Figure 6.5 Uncertainty reduction for annually (top) and seasonally aggregated (bottom 3 panels) national biosphere-atmosphere exchange for EU28 countries using the current ICOS (ICOS1) and future (ICOS2) network. Note the different colors for each point denote the number of atmospheric stations within a given country. Countries are ordered according to their area as represented within the model domain.

6.3.1.3 What do we learn from the future ICOS network?

Similar to section 6.3.1.1 for the ICOS1 network, we calculated the uncertainty reduction at grid scale also for the future ICOS2 network (Figure 6.3). The average uncertainty reduction for land regions is 17.4% and 11.8% for TM3-STILT and VUA, respectively. The addition of 33 stations relative to ICOS1, provide an additional 5% reduction in uncertainty according to TM3-STILT for pixel-resolved fluxes at seasonal time scales. VUA system shows a limited sensitivity to the different station configuration and the uncertainty reduction is only increased by 1.6%. As ICOS1 is already mainly located in central Europe, the expansion of the atmospheric network within the same area results in a limited increase of the uncertainty reduction. This can also be

seen by comparing Figure 6.2 and Figure 6.3 which unveils information on the spatial coverage achieved from the two ICOS networks. Both inversion systems show a fair information gain for central and west Europe even for ICOS1 with the limited number of stations. It is clear from the spatial distribution that Southern and Eastern Europe as well as the Iberian Peninsula do not benefit from the network expansion.

At country scale the network expansion significantly increases the average uncertainty reduction from 46% to 55% for the TM3-STILT system and from 32% to 36 % for the VUA. Well-constrained countries such as Germany and France share a quite large annual uncertainty reduction of 85%. The corresponding values for France are 85% and 79% for summer and fall, respectively, and 82% for Germany for both seasons. Those uncertainty reduction scores can be directly compared and are fully in line with the corresponding scores for the ICOS66 network in Kadygrov et al. (2015) study.

The limitation in uncertainty reduction in "station saturated" regions becomes more obvious when country scale is of interest. For example Austria with no inland stations, already has an annual uncertainty reduction of 75% and 39% in ICOS1, which in ICOS2 increases only to 78% and 41% for TM3-STILT and VUA systems respectively. Contrastingly Denmark with 1 station in ICOS1 and 2 stations in ICOS2 network has a drastic increase in uncertainty reduction from 36% and 34% to 52% and 43% for TM3-STILT and VUA respectively. This indicates that a future expansion of the network to regions that are currently insufficiently covered would be much more beneficial.

6.3.1.4 Impact of a network gap

To assess the impact of a spatial gap in the network, uncertainty reduction was computed also for a network with all stations missing in Germany. The country selection was made such that the gap would be well centered relative to the rest of the stations. Figure 6.6 shows the differences in uncertainty reduction between ICOS2 and ICOS2EG for all inversion systems. At pixel scale the uncertainty reduction is significantly reduced for the area of Germany by 15% to 35% for TM3-STILT system. The impact is weaker (around 5% to 10%) and less localized in the VUA inversion, in line with the previously mentioned more localized uncertainty reduction in the vicinity of observing stations in TM3-STILT.

At national scale results from the STILT-TM3 system show that the gap over Germany has direct impact on uncertainty reduction for Germany with a decrease of around 10%. The fact that the uncertainty reduction does not drop to smaller values is related to the impact of the network's strong coverage from stations in countries around Germany. On the other hand, neighboring countries show a decrease in the uncertainty reduction: Netherlands and Austria show a decrease of 1% to 3%, and Belgium, which has no inland stations, is the most affected country with a 12% reduction in uncertainty reduction compared to ICOS2 configuration. France, with 10 stations in ICOS2 configuration is also affected by the network gap showing a decrease of 5%. A slightly opposite impact is found for Spain (increase by 1%), which is an indication of uncertainties due to a limited ensemble size in the Monte Carlo method applied to retrieve the posterior uncertainties. For the VUA system the network gap resulted in similar impact but only for the directly affected area. The decrease in uncertainty reduction was found to be 11% for Germany, whilst for neighboring regions like Austria and France is limited to 1%.

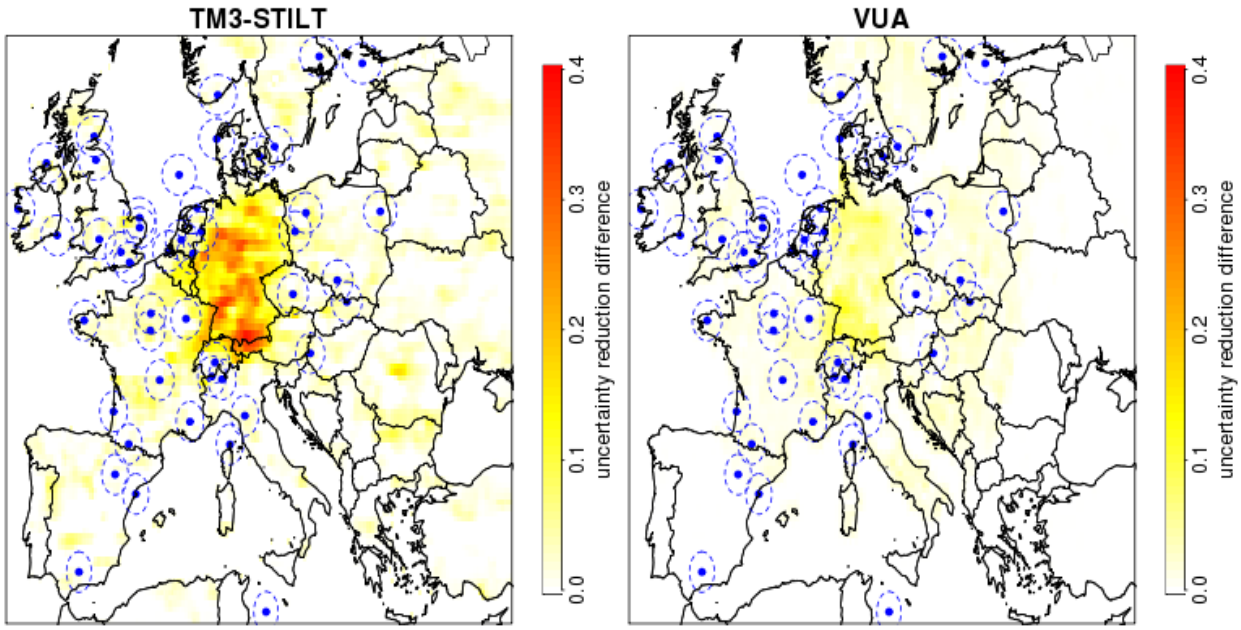

Figure 6.6 Difference in the annual uncertainty reduction between ICOS2 and ICOS2EG; the latter has a gap in the network for Germany.

6.3.2 How do biased observations impact flux retrievals?

In order to quantitatively assess the impact of measurement biases (in concentrations) and how that propagates to the fluxes, an uncorrelated measurement bias of 0.1 ppm was assumed. Spatially resolved impacts on retrieved fluxes are shown in Figure 6.7. The impact of the measurement bias shows to be of local nature and is limited in the vicinity of the station. Nevertheless the same bias seems to have an increased impact on the retrieved fluxes when assuming smaller model-data mismatches (Figure 6.7 subplots a, c, e, g). In detail, the averaged near field impact (averaging grid cells within 100 km of each site) were found to be 18.2, 23.6, 25.5, and 25.8 $gCm^{-2}y^{-1}$ for model-data mismatch error factors of 1, 0.3, 0.1 and 0.03. The near field distance value of 100 km is not arbitrary selected, but represents the spatial e-folding correlation scale assumed in the prior flux uncertainties. The impact seems to saturate for the smaller error factors.

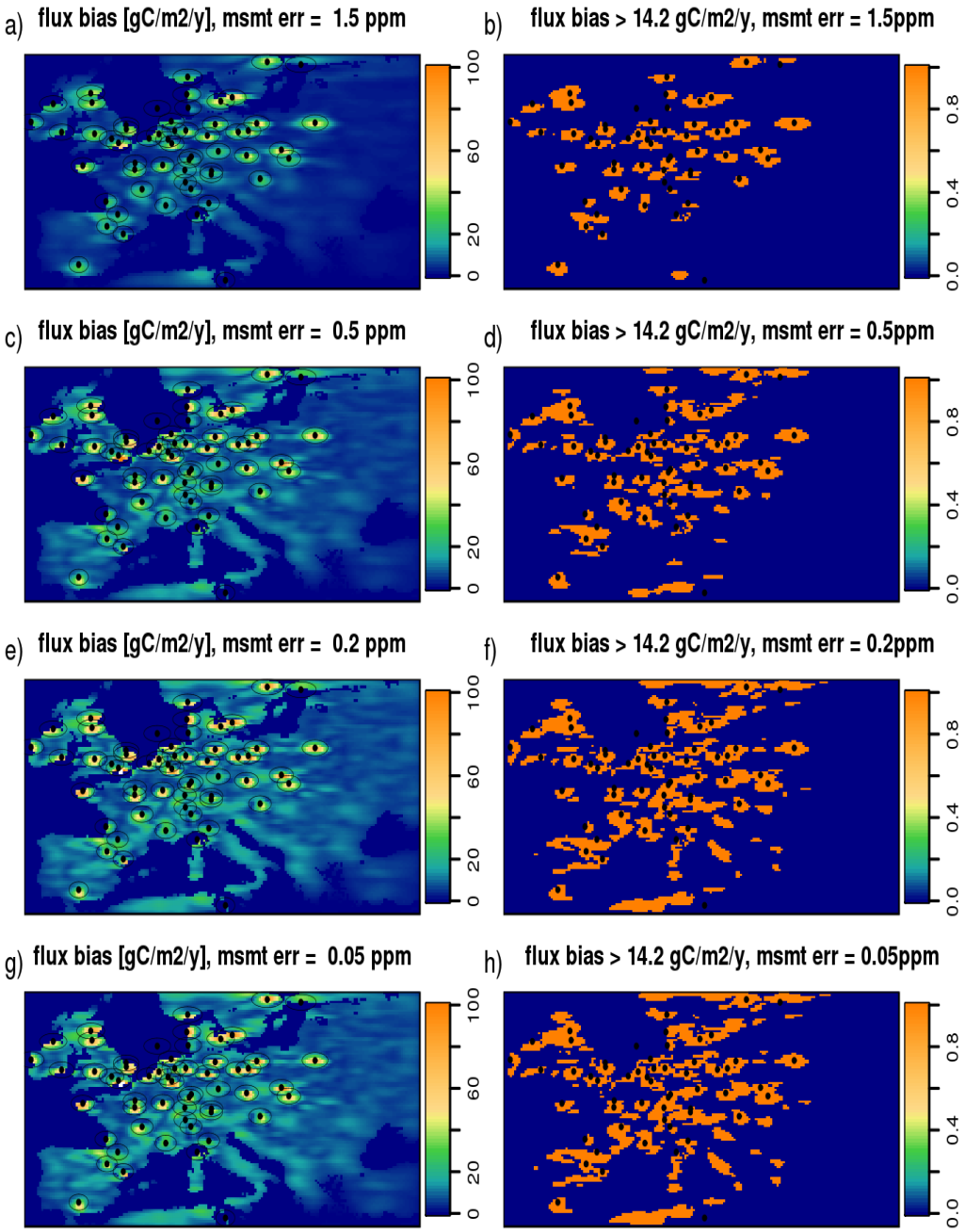

Figure 6.7 Impact on flux bias as calculated from eq.7 (left column) and land areas exceeding the threshold value of 14.2 $gCm^{-2}y^{-1}$ (right column) from 0.1 ppm measurement bias at each station and for different model-data mismatch errors (different rows). Units are given in $gCm^{-2}y^{-1}$.

At grid scale the targeted flux bias was selected such that allows resolving for interannual flux variations; thus was assigned with a threshold value of 14.2 $gCm^{-2}y^{-1}$. The second column in Figure 6.7 presents the areas that exceed this value. For the current transport models capabilities (Figure 6.7 subplot b) 8.4% of the land pixels were found not to comply with the threshold value. This fraction increases for future models with smaller transport errors with values of 12.4%, 16.1% and 16.8% for model-data mismatch error factors of 0.3, 0.1 and 0.03 respectively. Again

we see that when reducing model-data errors by a factor 10 or more compared to the current value, changes in the fraction of the area violating the threshold value seem to level off.

The retrieved flux biases at national and domain-wide aggregated scales, expressed as fractions of the respective prior uncertainties, are shown in Figure 6.8. Assuming a present network (error factor = 1, denoted with the black line), and comparing to the targeted 25% (posterior/prior uncertainty), flux biases are about a factor of 4 to 15 smaller for the different countries. For a future network where we expect smaller model-data mismatch (e.g. error factor = 0.03 denoted with the blue line), resulting flux biases are about a factor of 2.5 to 12 smaller than the threshold of 25% of the prior uncertainty. This indicates that a potential increase of the current bias requirement (0.1 ppm) in the atmospheric observations would not deteriorate the retrieved fluxes at national scale beyond the statistical uncertainty.

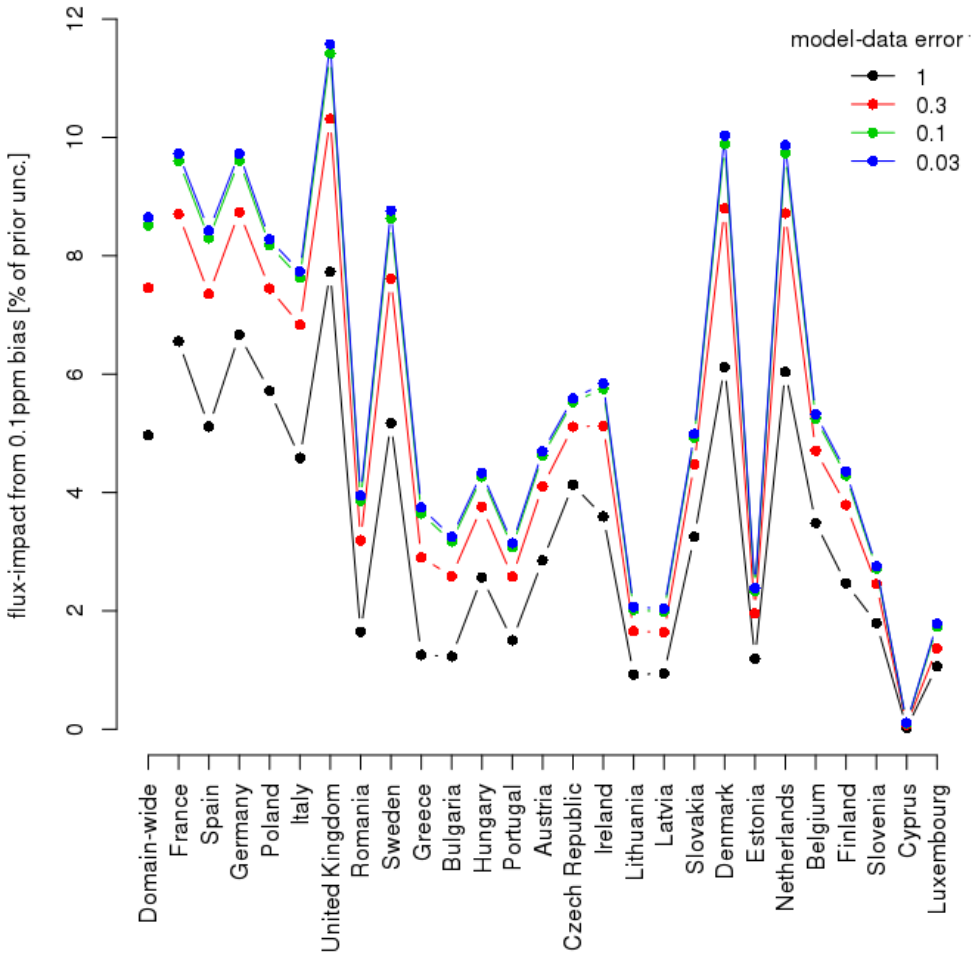

Figure 6.8 Impact of 0.1 ppm CO_2 measurement bias on retrieved fluxes for the domain-wide land area and for each EU28 countries, expressed as a fraction of a priori uncertainty in percent for different assumed model-data mismatch errors. Malta is omitted due to its small areal coverage. Countries are ordered by area.

6.4 Discussion

6.4.1 Uncertainty reduction metric

A coordinated network design study was made with two different regional inversion systems to estimate the uncertainty reduction at different spatial and temporal scales. By definition the uncertainty reduction depends on the prior and posterior uncertainties. Realistic uncertainties are a prerequisite when using uncertainty reduction as a metric for network design studies. For example using unrealistically large spatial correlation scales in prior uncertainties will result in a relatively large uncertainty reduction far away from the station location. This apparent knowledge in these regions though would be of artificial nature, since the stations do not directly "see" this region. As a consequence, the uncertainty reduction of a particular network configuration would be wrongly interpreted and potentially would result in a non-optimal network.

To provide a meaningful and realistic network assessment, this study utilizes an a priori error structure based on a direct comparison, of different a priori fluxes with eddy covariance observations (Kountouris et al., 2015). We argue that despite the resulting small spatial error correlation scales they can be regarded as realistic. Furthermore, to reveal potential patterns in the uncertainty reduction due to the transport uncertainties, we make use of different inversion systems. In such studies were multiple inversion systems are to be used a strict protocol should be considered. This is essential to harmonize the uncertainties for model-data mismatches and for prior fluxes but also to ensure the comparability of results by assuming common times and locations of observations between the inversion systems.

Extensive differences of the spatial distribution of the uncertainty reduction between the different inversion systems were found at seasonal and annual scales. Those differences are partially caused, by the different transport models involved that control, through their parameterization of vertical mixing (cloud venting, turbulent mixing in the boundary layer), the sensitivity of the network to upstream fluxes. Another cause for those seasonal differences, might be the slightly different treatment of the long-term bias (LTB) component, which has seasonal flexibility in the VUA system, but not in TM3-STILT. Note that both systems use a comparable resolution for the state space, the same prior uncertainty for domain-wide integrated fluxes, and the same uncertainty for model-data mismatch. Such large differences in the reduction of uncertainty could also be related to either differences in the transport model deployed, or to the specifics of the inversion systems optimization schemes. With respect to the transport, higher mixing heights are assumed for the VUA system, with the differences reaching values of 17%, 20%, and 26% for the MAM, JJA, and SON seasons compared to TM3-STILT. The error reduction directly depends on the footprint (transport). The latter is defined as the density of particles at a location $(\boldsymbol{x},t)$, that were transported backwards from the receptor location $(\boldsymbol{x}_r,t_r)$, normalized by the total number of particles released N_{tot}. The larger mixing high for the VUA system, results to smaller particle density, and consequently to a weaker footprint, hence lower error reduction. However this might not be the only mechanism which drives the observed differences in the uncertainty reduction, but only partially explains it. With regard to the optimization scheme, the EKF method deployed in the VUA system does show a slight artifact: Due to the limited number of members in the ensemble, the intended statistics is only imperfectly approximated. This has as a consequence that sometimes a change in a flux may correspond to an improvement of an "observed" concentration by coincidence instead of by causation. This so-called "localization

problem" causes the inversion system to somewhat exaggerate the error reduction in the long term. The problem is especially important with a large number of measurement sites, as is used here, because this causes an accumulation of small spurious contributions to error reduction. As already mentioned attempts were taken to suppress this "spurious reduction" (see also 6.2.2.2 section) which they unintended, also suppress a part of the real error reduction as no separation is possible.

The absolute level of uncertainty reduction is to a strong degree dependent on the prior uncertainties. Typical uncertainty reductions at seasonal scale are around 40-50% (TM3-STILT) at the close proximity of the atmospheric sites and for a pixel resolution of 0.25°. At the national scale, values around 40% to 60% were found. This relatively small reduction is related to the tight prior error constraint (0.3 GtC y-1 annual domain-wide integrated uncertainty).

The dependence of uncertainty reduction on the distance to the nearest station (Figure 6.4) shows less of a decay for the future ICOS2 network compared to ICOS1 network, which can be attributed to the increased station density resulting in an increased uncertainty reduction at regions in between station locations.

Whilst only minor seasonal differences were identified, it is interesting to note that seasonal variations at country scale are decreasing for regions which are strongly influenced by the denser ICOS2 network compared to ICOS1. For example seasonal variations (standard deviation of national uncertainty reduction across the three seasons) between ICOS1 and ICOS2 networks for Denmark, Germany and United Kingdom were decreased by 31%, 53% and 65% respectively. Contrastingly countries that have minor benefits from ICOS2 expansion and they already lacking stations in ICOS1, do have similar or even increased seasonal variations for ICOS2 (e.g. Spain, Portugal and Greece have an increase in seasonal variations of 58%, 25% and 26% respectively). This indicates that regions not well constrained by the atmospheric network are more sensitive to the atmospheric transport. In fact, seasonal differences in mixing height for the TM3-STILT scheme were found to be 15%, 11%, and 26% relative to the seasonally averaged mixing height.

6.4.2 Measurement bias impact on flux retrievals

The assessment of the impact of 0.1 ppm uncorrelated measurement bias on retrieved fluxes was made using the Jena TM3-STILT regional inversion system for the ICOS2 network. Model-data mismatch was set to reflect current inverse transport model capabilities, but also future systems with reduced model-data mismatch (up to a factor 30 smaller than current models).

The study indicates that when one is interested in fluxes at smaller spatial scales, measurement requirements are tighter than when solving for large-scale fluxes. The results show that measurement biases of 0.1 ppm for CO_2 are hardly compatible with uncertainty targets for retrieved spatially resolved fluxes that require flux signals to be resolved on annual-interannual time scales. Larger measurement biases will lead to retrieval biases that easily become much larger than the expected variations in fluxes. So relaxing those requirements will reduce the capability to detect changes in fluxes. We note that the assumed interannual variability of 71 $gCm^{-2}y^{-1}$ might be erroneous at some extent as it comes only from model intercomparisons and might not coincide with the one measured with eddy covariance flux towers. As this information is not available and it would also be beyond the scope of this study, the current estimate can be considered as representative.

The average flux bias within a radius of 100 km from a station was found to be 18.17 $gCm^{-2}y^{-1}$ which results in a total flux bias for this area of 570 $ktCy^{-1}$. Comparing with the statistically derived uncertainties for the same areas within 100 km distance from the stations, this flux bias is significant. Specifically, the averaged uncertainty reduction for these areas is around 36%, and given the prior uncertainty (0.3 $GtCy^{-1}$ domain wide, homogeneously distributed according to the land area), the posterior uncertainty equals to 600 $ktCy^{-1}$. To conclude, both metrics, i.e. the comparison of flux bias to the posterior uncertainty, and the comparison of the flux bias to the targeted uncertainties for resolving fluxes at interannual scale, leaves no room for relaxing the measurement accuracy. Of note is that smaller or even zero biases occur when moving away from a given station, especially for Eastern Europe. In such areas where there are no measurements, there is also no information gain.

For domain-wide or national aggregated annual fluxes the uncertainty reduction achieved by current inversions (about 75% for CO_2 inversions) is not significantly impacted by measurement biases. For such aggregated scales the bias impacts retrieved fluxes when it exceeds about 0.4 ppm.

It is worth to mention the dependence of the results on the error structure. This is due to the nature of the Bayesian inversions, since the results depend on the exact settings of the prior uncertainties including the magnitude, and the error structure in terms of spatial and temporal correlations. In the current study we do not assess this dependence, but use a data driven error structure with a constant in time uncertainty and rather small spatial correlations (around 100 km) as those calculated in Kountouris et al. (2015) study. The question that immediately rises is what would have been the impact in the flux biases, if larger spatial correlation lengths would have been assumed, as this is the case for many inversion systems. In Rödenbeck et al. (2006) study, the assumed spatial correlation was in the order of thousand km, much larger than our assumption. Unfortunately comparisons and safe conclusions can not be drawn due to substantial differences between the two studies. In Rödenbeck et al. (2006) a global network was used but biases in observations were introduced to only 5 stations. On the contrary we are focused on regional scale and we introduce biases in the full network. It is clear that meaningful flux biases can be retrieved only within regions that are influencing the atmospheric stations. Nevertheless flux biases for all TransCom 3 regions and for roughly twice as large measurement bias assumption [2], did not exceed values of 0.1 $PgCy^{-1}$. In our study we assumed 0.1 ppm measurement bias and we calculated a resulting flux bias of 0.04 $PgCy^{-1}$.

We note that this study uses a constant measurement bias as this is proposed by the WMO. Although the aim is to infer if the current WMO policy on atmospheric measurements is enough to retrieve meaningful flux estimates, it would be very interesting to investigate in depth the error structure of the measurement bias at site level. This would demand detailed information from the different institutions on whether the bias is constant over decades or changes whenever a calibration gas is exchanged. This information could be added to the model – data mismatch error in the inversion systems making them less susceptible to biased retrievals. An attempt to correct biased measurements was made from Corazza et al. (2011), by introducing site specific biases for their N_2O inversions.

[2] Biases were calculated as the difference between flask and continuous measurements at coincidental times. The residual time series then are low-pass filtered on a time scale of one year resulting to annually constant observation bias of approximately 0.2 ppm.

The model-data error assigned in the observation error covariance matrix depends on the station type. The uncertainties assumed (e.g. 1.5 ppm for tall towers) are pertained to weekly measurements, which are inflated to hourly. The error inflation is required to assure a balanced impact when flask (normally weekly) and continuous (normally hourly) measurements are combined (see also Rödenbeck (2005)). At annual scale the resulted uncertainty, defined as the ratio of the weekly uncertainty divided by the square root of the number of the weeks within a year, would be much smaller with a value of 0.2 ppm. Kountouris et al. (2016b) discussed how sensitive fluxes are, in the presence of biased fossil fuel flux fields. The model-data error should also reflect uncertainties in the fossil fuel signal although it is traditionally assumed to be perfectly known by the inversion community. We note that the annual uncertainty in the current study might be smaller than the mismatch in true and assumed fossil fuel signal at some of the stations. Peylin et al. (2011) estimated the annual averaged standard deviation of hourly averaged fossil fuel CO_2 concentrations. They used a number of different emission inventories and they performed forward runs with the same transport model. They calculated for SCH and CMN stations an uncertainty in the fossil fuel component of 0.35 and 0.23 ppm respectively which is already larger than the annual uncertainty assumed in the current study.

Chapter 7

Synthesis of the results – Outlook

The chapters of this thesis follow a logical and sequential approach, which aims to quantitatively address a network design study for the ICOS atmospheric network. As already mentioned, the network design study uses an atmospheric inversion system based on the Bayesian approach, which requires a-priori information and a well defined prior error structure. The latter determines the spread of the information for a given network, controlled by the assumed spatial and temporal correlations. Afterwards, the error structure should be implemented in the inversion system and subsequently the system should be extensively tested. Following, the inversion system can be applied to real measurements and its performance can be assessed. Only after the completion of all those preparatory steps a network design study can be conducted.

The first part of the thesis (chapter 3) is dedicated to prior flux uncertainties, and investigates the error structure of flux residuals for correlations in time and in space. Next, this information is used, to build a data driven error covariance matrix which is needed in atmospheric inversions. In the second part of this thesis (chapter 4), the performance of the inverse modeling system at regional scale is explored. The aim of this assessment is to investigate to what extent, with respect to temporal and spatial scales, we can constrain regional carbon fluxes. Further, the European carbon budget and carbon fluxes are addressed at high spatial resolution and validated using eddy covariance flux measurements (chapter 5). In the third and last part (chapter 6), this thesis demonstrates a network assessment in terms of the uncertainty reduction at annual and seasonal temporal scales and domain-wide down to grid-cell level for the spatial scales. Moreover, it quantifies flux biases in the presence of biases in atmospheric measurements, by propagating this bias to the flux space.

This chapter summarizes the research performed in this thesis and aggregates the conclusions to answer the research questions as they are formulated in the introduction:

1. What is the error structure of the prior error covariance used in atmospheric CO_2 inversions? Can we generalize and use this structure in every inversion system?

2. What can we gain from, and what are the limitations of a regional high resolution Inversion system? What is the biospheric carbon budget in Europe estimated using data driven prior uncertainties and a regional high resolution inversion system?

3. How well does the current ICOS network perform, and what will be gained from the future expansion? Does a bias in atmospheric observations affect flux retrievals, and how accurate should observations be?

7.1 On the prior error

7.1.1 What is the error structure of the prior error covariance used in atmospheric CO_2 inversions?

The first question addresses the prior error covariance structure. Spatial and temporal autocorrelations of model-data flux residuals can be approximated reasonably well with an exponentially decaying model. It is relatively easy to implement while using such a model in the inversion systems; it also keeps computational costs low, and ensures a positive definite covariance matrix.

Weak spatial correlations were found and they are limited to lengths up to 40 km. This finding is supported also by the autocorrelation analysis of aircraft measurements, which solves the spatial flux variability at smaller scales than the inferred correlations using eddy covariance data. Spatial e-folding correlations for model-model residuals were found to be significantly larger ranging from 260 km to 1000 km approximately. This diference can mainly be assigned due to features that the biosphere models share. For instance, temperature, downward radiation, and meteorological fields have significant commonalities, resulting to an increased similarity between the modeled fluxes.

Autocorrelation e-folding times for residuals between modeled and observed fluxes were found to have a range of around 26 to 70 days. This difference is diminished when excluding sites with large residual bias, while autocorrelation times become coherent among all three biosphere models with a value of around 30 days. Resulted autocorrelation times from model-model flux residuals confirmed also the temporal scale found in model-data analysis for the majority of the ensemble members.

7.1.2 Can we generalize and use this error structure in every inversion system?

A fully defined and quantified error structure, which can be applied in every atmospheric inversion system, regardless the system itself and the prior that uses, would be highly attractive. Bringing together results from both model-data and model-model analysis, it seems to be reasonable to generalize and use the same temporal error correlation in atmospheric inversions. The recommended autocorrelation time is around 30 days.

Spatial correlations did not converge between model-data and model-model residuals. To generalize the spatial error structure, a careful assessment needs to be held. On the one hand, all models showed very short spatial e-folding lengths in the model-data residual analysis, but for a representative resolution of around 1 km. On the other hand, model-model analysis showed larger e-folding lengths for coarser spatial flux resolutions up to 50 km x 50 km. Considering that inversion systems have usually a state space much coarser than 1 km, it is reasonable to assume a correlation length which lies between the model-data and model-model calculations. In any case, this research emphatically excludes correlations of several hundreds of km assumed in many inversion studies. A reasonable correlation length would be around 100 km.

We note that although temporal and spatial correlation lengths can be generalized, the variance of model-data residuals, which describes the main diagonal of the prior error covariance matrix, needs to be defined explicitly depending on the prior used by the inversions.

Assuming such small scale correlations in the prior error covariance matrix, aggregating the uncertainty over larger spatial and temporal scales results to very small prior uncertainties. For example, assuming correlations of 100 km and 30 days with diagonal elements of 2.27 μmoles m^{-2} s^{-1}, the aggregated uncertainty over the current domain of interest, and over the full year, counts only for 0.15 $GtCy^{-1}$. Although there is no concrete or correct method to quantify precisely what the correct uncertainty budget would be, this prior uncertainty seems to be unrealistically small and error inflation should take place. Many studies are inflating the error covariance matrix by scaling it accordingly. This thesis, instead of the traditional error inflation, implements and suggests a more formal way to inflate the uncertainty by introducing a global bias term presented in chapter 4.

7.2 On the regional atmospheric inversions

7.2.1 What can we gain from and what are the limitations of a regional high resolution inversion system?

The system performance was carefully assessed at several spatial and temporal scales in both flux and observation space. Results from the synthetic inversions seem to be very promising. The spatiotemporal flux variability was fully captured within the statistical uncertainties, concerning the domain-wide down to country and the annual, down to monthly scales. Moreover, the inversion system was able to significantly correct monthly fluxes at grid scale, representing a hypothetical validation using eddy covariance flux data. Additionally, we showed that eddy covariance measurements can be used principally for validation of the inverse estimates at monthly timescales. This good performance was achieved although two strict rules were adopted: 1) prior and known truth was produced from different models and 2) the network had the same number of observations and data gaps as in the real data inversion. This gave us confidence that such an atmospheric inverse system, using observed CO_2 dry mole fractions from the current European network, is capable to constrain terrestrial fluxes at least for the abovementioned scales. After that, analysis was conducted so as to reveal the temporal and spatial scales at which the atmospheric inversion performs best. We discovered that the system maximizes the information gain at 200 km spatial and at seasonal temporal scales.

7.2.2 What is the biospheric carbon budget in Europe using data driven prior uncertainties and a regional high resolution inversion system?

The calculated annual carbon budget for 2007, considering only biospheric fluxes and using a state of the art modeling scheme at high resolution (Jena inversion system), is calculated to be between 0.23 and 0.38 GtC y^{-1} depending on the assumed error structures. Extensive sensitivity tests were held to provide robust regional carbon estimates. The impact of different network configurations (to investigate a potential misrepresentation of the nocturnal boundary layer), as

well as fossil fuel scenarios, were assumed and investigated. Regional scale flux estimations are in agreement within the uncertainty range, with results derived from global systems. Cross validation of the flux retrievals against eddy covariance flux measurements at local scale showed a significant correction. It is worth to highlight that the agreement between retrieved flux estimates and observed fluxes for non-crop sites found to be remarkably good. Contrastingly, crop sites showed the largest mismatches, which are probably linked to the high human interaction in such agricultural areas.

Five different error structures were applied in the modeling scheme to assess their impact on the retrieved fluxes. The error structures can be classified into two distinct categories, using or not one more degree of freedom, allowing for a bias correction. The error structure, without the bias term, is inflated to the targeted annual domain-wide integrated uncertainty, whilst the inflation for the other error structure is introduced by the bias term. The estimated carbon budgets derived from the different error structures do not significantly differ, as they lie within the posterior uncertainties.

7.3 On the network design

7.3.1 How well does the current ICOS network perform, and what will be the gain due to future expansion?

Currently, the ICOS network comprises about 25 stations, mainly located in central Europe. Future plans anticipate for an ambitious network of more than 50 stations. With the current network fluxes on land pixels show on average an annual uncertainty reduction of 12.5% for TM3-STILT, and 10.2% for the VUA system. At country scale clearly, countries in central and Western Europe benefit from the denser network, and show an uncertainty reduction with typical values of 60% – 70% for TM3-STILT. When the future ICOS network will be established, an additional reduction in uncertainty of 5% may be expected for pixel-resolved fluxes at seasonal time scales. Moreover, the average uncertainty reduction at country scale is expected to increase from 46% to 55% (TM3-STILT system). Nevertheless, we note a limitation in the information we gain for “station saturated” regions; this indicates that a future expansion to under-constrained regions would be much more beneficial.

The uncertainty reduction study revealed also valuable information regarding the information gain we could expect from a given station. The average relative uncertainty reduction over the stations is decreased by 50%, with respect to its initial value within the first 500 km. This localized nature of the uncertainty reduction confirms the need for establishing a dense atmospheric network. It also confirms that retrieved flux estimates are expected to be highly uncertain for the poorly constrained Eastern and Southern European regions.

7.3.2 Does a bias in atmospheric observations affect flux retrievals, and how accurate should observations be?

In order to answer the question, if biases in observations affect significantly the retrieved flux estimates, firstly, we must formulate the scientific question, and secondly, we must decide the scale of interest. Depending on the target we want to set, the impact could be totally different.

For example, we are interested to resolve fluxes at grid scale, and further, we would like to detect flux variations at interannual temporal scales. A measurement bias of 0.1 ppm, as suggested by WMO, seems to be the upper limit, and there is no room for relaxing the measurement requirements. Contrastingly, if we are interested in country scale annually aggregated fluxes, then the measurement bias could be relaxed, although it is not recommended. Of note is though, that future inversion systems are expected to have reduced transport errors, and this might increase the impact of measurement biases.

7.4 Outlook

This study tried to compose a synthetic experiment in the most realistic way, to evaluate flux retrievals derived from the Jena inversion system. Notwithstanding, there is still space to design a more realistic and challenging experiment by producing pseudo data, using a different transport model, driven with different meteorological fields than the ones used in the inverse procedure. In such a case, this would allow to further investigation of transport uncertainties in the experiment and explore the inversion limitations more accurately.

The atmospheric network sparseness was made clear throughout this study, especially for the Eastern and Southern European parts. It is a necessity to incorporate more data streams into the inversion schemes to constrain more regions. To that direction satellite observations can be proved very helpful, as they can be provided with a dense spatial coverage. The first inversions using space-borne measurements are already a reality show, however, large differences compared to inversions based on data from ground stations. Combined data streams coming from satellites and ground observations can, and should be extensively tested and validated in the future regional inversion studies.

More research is needed to investigate the impact on flux retrievals in the presence of correlated measurement biases. In the current study (see also chapter 6) biases were assumed to be uncorrelated from station to station. However, we must consider the fact that correlations might not only be present, but also they are institution or country specific.

One important source of uncertainty in atmospheric inverse modeling is the mixing height in the transport models. More specific, the mixing height is a key parameter needed to determine the turbulent domain, in which dispersion takes place, and it affects the tracer concentration directly. Several methods have been developed for the estimation of mixing heights using in-situ data obtained e.g. with the use of radiosondes and aircrafts, but those methods are rather expensive, while the data do not have a satisfactory spatial and/or temporal coverage. An alternative to those measurements are remote sensing techniques like LIDARs and Ceilometers, with the advantage of the continuous observation of the planetary boundary layer (PBL). In future, the German weather service (DWD) is planning to operationally derive mixing height data, from a very dense network of ceilometers (around 150) across Germany. The derived mixing height information will draw a lot of interest, when it will be evaluated and incorporated in the regional atmospheric inversions. This would allow potentially, retrieving carbon fluxes with reduced uncertainties, by incorporating more accurate mixing height estimations.

Bibliography

Albergel, C., Calvet, J.-C., Gibelin, A.-L., Lafont, S., Roujean, J.-L., Berne, C., Traullé, O., and Fritz, N.: Observed and modelled ecosystem respiration and gross primary production of a grassland in southwestern France, Biogeosciences, 7, 1657-1668, doi:10.5194/bg-7-1657-2010, 2010.

Alemanno, M., Di Diodato, A., Lauria, L. and Santobuono, N.: Environmental measurements at Monte Cimone GAW station, Int. J. Global Warming, Vol. 6, No. 4, 424–454, doi: 10.1504/IJGW.2014.066048 2014.

Allard, V., Ourcival, J.-M., Rambal, S., Joffre, R., and Rocheteau, A.: Seasonal and annual variation of carbon exchange in an evergreen Mediterranean forest in southern France, Global Change Biology, 14(4), 714-725, 2008.

Ammann, C., Spirig, C., Leifeld, J., and Neftel, A.: Assessment of the nitrogen and carbon budget of two managed grassland fields, Agriculture, Ecosystems and Environment, 133, 150–162, 2009.

Aubinet, M., Grelle, A., Ibrom A., Rannik Ü., Moncrieff J., Foken T., Kowalski A.-S., Martin P.-H., Berigier P., Bernhofer C., Clement R., Elbers J., Granier A., Grünwald T., Morgenstern K., Pilegaard K., Rebmann C., Snijders W., Valentini, R. and Vesala, T.: Estimates of the Annual Net Carbon and Water Exchange of Forests: The EUROFLUX Methodology, Adv. Ecol. Res. 30 , 113-175, 2000.

Aubinet, M., Chermanne, B., et al.,: Long term carbon dioxide exchange above a mixed forest in the Belgian Ardennes, Agricultural and Forest Meteorology 108(4): 293-315, 2001.

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

Baldocchi, D., et al., FLUXNET: A new tool to study the temporal and spatial variability of ecosystem–scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82, 2415–2434, 2001.

Barcza, Z., Weidinger, T., Csintalan., Zs., Dinh, N.-Q., Grosz, B., Tuba, Z.: The carbon budget of a semiarid grassland in a wet and a dry year in Hungary, AGR ECOSYST ENVIRON 121: (1-2)21-29, 2007.

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, Atmos. Chem. Phys., 13, 8695–8717, doi:10.5194/acp-13-8695-2013, 2013.

Bates, D.-M. and Watts, D.-G.: Nonlinear Regression Analysis and Its Applications, Wiley, 1988

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M. and Monfray, P.: Inverse modeling of annual atmospheric CO_2 sources and sinks: 1. Method and control inversion, Journal of Geophysical Research: Atmospheres (1984--2012) 104, 26161-26178, 1999.

BP (British Petroleum): Statistical Review of World Energy 2012: http: //www.bp.com/statisticalreview, last access: December 2013, 2012.

Broquet, G., Chevallier, F., Bréon, F. M., Kadygrov, N., Alemanno, M., Apadula, F., Hammer, S., Haszpra, L., Meinhardt, F., Morguí, J. A., Necki, J., Piacentino, S., Ramonet, M., Schmidt, M., Thompson, R. L., Vermeulen, A. T., Yver, C., and Ciais, P.: Regional inversion of CO_2 ecosystem fluxes from atmospheric measurements: reliability of the uncertainty estimates, Atmos. Chem. Phys., 13, 9039-9056, doi:10.5194/acp-13-9039-2013, 2013.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T.and Ciais, P.: A European summertime CO_2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, J Geophys. Res., 116, D23303, doi:10.1029/2011JD016202, 2011.

Carouge, C., Bousquet, P., Peylin, P., Rayner, P. J., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 1: Potential of the 2001 network, Atmos. Chem. Phys., 10, 3107-3117, doi:10.5194/acp-10-3107-2010, 2010a.

Carouge, C., Peylin, P., Rayner, P. J., Bousquet, P., Chevallier, F., and Ciais, P.: What can we learn from European continuous atmospheric CO_2 measurements to quantify regional fluxes – Part 2: Sensitivity of flux accuracy to inverse setup, Atmos. Chem. Phys., 10, 3119-3129, doi:10.5194/acp-10-3119-2010, 2010b.

Casals, P., Lopez-Sangil, L., Carrara., A., Gimeno, C., and Nogues, S.: Autotrophic and heterotrophic contributions to short-term soil CO_2 efflux following simulated summer precipitation pulses in a Mediterranean dehesa, Global Biogeochemical Cycles, 25 (3), doi: 10.1029/2010GB003973, 2011.

Chevallier, F., Palmer, P. I., Feng, L., Bösch, H., O'Dell, C., and Bousquet, P.: Towards robust and consistent regional CO_2 flux estimates from in situ and space-borne measurements of atmospheric CO_2, Geophys. Res. Lett., 41, 1065–1070, doi:10.1002/2013GL058772, 2014.

Chevallier, F., Wang T., Ciais P., Maignan F., Bocquet M., Altaf A. M., Cescatti A., Chen J., Dolman A., J., Law B. E., Margolis, H. A., Montagnani, L., Moors, E. J.: What eddy-covariance measurements tell us about prior land flux errors in CO_2 flux inversion schemes, Glob. Biogeochem. Cy., 26, GB1021, doi:10.1029/2010GB003974, 2012.

Chevallier F., Viovy N., Reichstein M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO_2 surface fluxes, doi: 10.1029/2006GL026496, 2006.

Chiesi M., Fibbi L., Genesio L., Gioli B., Magno R., Maselli F., Moriondo M., Vaccari F.,-P: Integration of ground and satellite data to model Mediterranean forest processes. International Journal of Applied Earth Observation and Geoinformation, 13, 504–515, doi: 10.1016/j.jag.2010.10.006, 2011.

Ciais, P., Paris, J. D., Marland, G., Peylin, P., Piao, S. L., Levins, I., Pregger, T., Scholz, Y., Friedrich, R., Rivier, L., Houwelling, S., Schuldze, E. D., and members of the CARBOEUROPE SYNTHESIS TEAM: The European carbon balance. Part1: fossil fuel emissions, Glob Change Biol, 16, 1395-1408, doi: 10.1111/j.1365-2486.2009.02098.x, 2009.

Ciais, P., Borges, A. V., Abril, G., Meybeck, M., Folberth, G., Hausglustaine, D., and Janssens, I. A.: The impact of lateral carbon fluxes on the European carbon balance, Biogeosciences, 5, 1259-1271, doi: 10.5194/bg-5-1259-2008, 2008.

Ciais, P., Peylin, P. and Bousquet, P.:, Regional biospheric carbon fluxes as inferred from atmospheric CO_2 measurements, Ecol Appl, 10, 1574-1589, doi: 10.2307/2641225, 2000

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, Q.J.R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828, 2011.

Delpierre, N., Soudani, K., François, C., Köstner, B., Pontailler, J.-Y., Nikinmaa, E., Misson, L., Aubinet, M., Bernhofer, C., Granier, A., Grünwald, T., Heinesch, B., Longdoz, B., Ourcival, J.-M., Rambal, S., Vesala, T., and Dufrêne, E.: Exceptional carbon uptake in European forests during the warm spring of 2007: a data-model analysis, Glob. Change Biol., 15, 1455–1474, 10 doi:10.1111/j.1365-2486.2008.01835.x, 2009.

Deng, F., Jones, D. B. A., Henze, D. K., Bousserez, N., Bowman, K. W., Fisher, J. B., Nassar, R., O'Dell, C., Wunch, D., Wennberg, P. O., Kort, E. A., Wofsy, S. C., Blumenstock, T., Deutscher, N. M., Griffith, D. W. T., Hase, F., Heikkinen, P., Sherlock, V., Strong, K., Sussmann, R., and Warneke, T.: Inferring regional sources and sinks of atmospheric CO_2 from GOSAT XCO_2 data, Atmos. Chem. Phys., 14, 3703–3727, doi:10.5194/acp-14-3703-2014, 2014.

Dietiker, D., Buchmann, N., Eugster, W.: Testing the ability of the DNDC model to predict CO_2 and water vapour fluxes of a Swiss cropland site, Agriculture, Ecosystems and Environment, 139, 396-401, 2010.

Dolman, A.-J., Noilhan, J., Durand, P., Sarrat, C., Brut, A., Piquet, B., Butet, A., Jarosz, N., Brunet, Y., Loustau, D., Lamaud, E., Tolk, L., Ronda, R., Miglietta, F., Gioli, B., Magliulo, V., Esposito, M., Gerbig, C., Korner, S., Glademard, R., Ramonet, M., Ciais, P., Neininger, B., Hutjes, R-W.-A., Elbers, J.-A., Macatangay, R., Schrems, O., Perez-Landa, G., Sanz, M.-J., Scholz, Y., Facon, G., Ceschia, E., Beziat, P.: The CarboEurope Regional Experiment Strategy. Bulletin of the American Meteorological Society, 87(10), 1367-1379. (2006).

Dlugokencky, E.J., Lang, P.M., Masarie, K.A., Crotwell, A.M., and Crotwell, M. J.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2014, Version: 2015-08-03, Path: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/.

Draxler, R. R. and Hess, G.: An overview of the hysplit 4 modelling system for trajectories. Australian Meteorological Magazine, 47(4), 1998.

Eidenshink, J. C. and Faundeen, J. L.: The 1 km AVHRR global land data set: first stages in implementation, Int. J. Remote Sens., 15(17), 3443–3462, 1994.

Enting, I. G., Trudinger, C. M., and Francey, R. J.: A synthesis inversion of the concentration and $\delta^{13}C$ of atmospheric CO_2, Tellus, ser.B, 47, 35-52, 1995.Ferrarese S., Apadula F., Bertiglia

F., Cassardo C., Ferrero A., Fialdini L., Francone C., Heltai D., Lanza A., Longhetto A., Manfrin M., Richiardone R., Vannini C.: Inspection of high–concentration CO2 events at the Plateau Rosa Alpine station, Atmospheric Pollution Research 6, 415-427, doi:10.5094/APR.2015.046, 2015.

Etzold, S., Buchmann, N., and Eugster, W.: Contribution of advection to the carbon budget measured by eddy covariance at a steep mountain slope forest in Switzerland, Biogeosciences, 7, 2461-2475, doi: 10.5194/bg-7-2461-2010, 2010.

Farquhar, G., G., von Caemmerer, S., Berry, J.-A.: A biochemical model of photosyntetic CO_2 assimilation in leaves of C_3 species. Planta 149, 78-90, 1980.

Feng, L., Palmer, P. I., Parker, R. J., Deutscher, N. M., Feist, D. G., Kivi, R., Morino, I., and Sussmann, R.: Estimates of European uptake of CO_2 inferred from GOSAT XCO_2 retrievals: sensitivity to measurement bias inside and outside Europe, Atmos. Chem. Phys., 16, 1289-1302, doi: 10.5194/acp-16-1289-2016, 2016.

Ferrarese S., Apadula F., Bertiglia F., Cassardo C., Ferrero A., Fialdini L., Francone C., Heltai D., Lanza A., Longhetto A., Manfrin M., Richiardone R., Vannini C.:. Inspection of high–concentration CO_2 events at the Plateau Rosa Alpine station, Atmos. Pollution Res., 6, 3, 415-427, doi:10.5094/APR.2015.046, 2015.

Fiedler, V., Dal Maso, M., Boy, M., Aufmhoff, H., Hoffmann, J., Schuck, T., Birmili, W., Hanke, M., Uecker, J., Arnold, F., and Kulmala, M.: The contribution of sulphuric acid to atmospheric particle formation and growth: a comparison between boundary layers in Northern and Central Europe, Atmos. Chem. Phys., 5, 1773-1785, doi:10.5194/acp-5-1773-2005, 2005.

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

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C. and Zeng, N.: Climate–Carbon Cycle Feedback Analysis: Results from the C4MIP Model Intercomparison. J. Climate, 19, 3337–3353. doi: http://dx.doi.org/10.1175/JCLI3800.1, 2006.

Friedl, M. A., McIvera, D. K., Hodgesa, J. C. F., Zhanga, X. Y., Muchoneyb, D., Strahlera, A. H., Woodcocka, C. E., Gopala, S., Schneidera, A., Coopera, A., Baccinia, A., Gaoa, F., and Schaafa, C.: Global land cover mapping from MODIS: algorithms and early results, Remote sensing of environment 83, 287-302, doi:10.1016/S0034-4257(02)00078-0, 2002.

Gabriel, P., Gielen, B., Zona, D., Rodrigues, A., Rambal, S., Janssens, I., Ceulemans, R.: Carbon and water vapor fluxes over four forests in two contrasting climatic zones, Agricultural and Forest Meteorology 180, 211– 224, 2013.

Garbulsky, M.-F., Penuelas, J., Papale, D., Filella, I.: Remote estimation of carbon dioxide

uptake by a Mediterranean forest, GLOBAL CHANGE BIOLOGY, vol. 14, 2860-2867, ISSN: 1354-1013, doi: 10.1111/j.1365-2486.2008.01684.x, 2008.

Garrigues, S., Olioso, A., Calvet, J-C., Martin, E., Lafont, S., Moulin, S., Chanzy, A., Marloie, O., Desfonds, V., Bertrand, N., Renard, D.: Evaluation of land surface model simulations of evapotranspiration over a 12 year crop succession: impact of the soil hydraulic properties, Hydrology and Earth System Sciences Discussion, 11, 11687-11733. doi:10.5194/hessd-11-11687-2014, www.hydrol-earth-syst-sci-discuss.net/11/11687/2014/, 2014.

Gerbig, C., Körner. S. and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmos. Chem. Phys., 8, 591-602, doi:10.5194/acp-8-591-2008, 2008.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S. and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms , J Geophys Res – Atmos, 108, 4756, doi: 10.1029/2002JD003018, 2003a.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bakwin, P. S. and Grainger, C. A.: Toward constraining regional-scale fluxes of CO_2 with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework , J Geophys Res, 108, 4757, doi: 10.1029/2003JD003770, 2003b.

Gielen, B., B De Vos, Campioli, M., Neirynck, J., Papale, D., Verstraeten, A., Ceulemans, R., Janssens, I.: Biometric and eddy covariance-based assessment of decadal carbon sequestration of a temperate Scots pine forest. Agricultural and Forest Meteorology 174, 135-143, 2013.

Gioli, B, Miglietta, F, Vaccari, F.-P, Zaldei, A., De Martino, B.: The Sky Arrow ERA, an innovative airborne platform to monitor mass, momentum and energy exchange of ecosystems. Annals of Geophysics, 49, n. 1, pp 109-116, 2006.

Göckede, M., Michalak, A. M., Vickers, D., Turner, D. P., and Law, B. E.: Atmospheric inverse modeling to constrain regional scale co2 budgets at high spatial and temporal resolution. J. Geophys. Res., 115(D15113), 2010.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, M.-J., Brunet, Z., et al.,: 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.

Groenendijk, M., Dolman A.-J., van der Molen, M.-K., Leunung, R., Arneth, A., Delpierre, N., Gash, J.-H.-C., Lindroth, A., Richardson, A.-D., Verbeeck, H., and Wohlfahrt., G.: Assessing parameter variability in a photosynthesis model within and between plant functional types using global Fluxnet eddy covariance data, Agr. Forest Meteorol. 151, 22-38 , doi: 10.1016/j, 2011.

Guidolotti, G., Rey, A., D'Andrea, E., Matteucci, G., De Angelis, P.: Effect of environmental variables and stand structure on ecosystem respiration components in a Mediterranean beech forest, Tree Physiology 33: 960-972 (doi:10.1093/treephys/tpt065), 2013.

Gurney, K. R., Rachel M. L., Denning, A. S., Rayner, P. J., Bernard C. P., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, Fung, I. Y., Heimann, M., John, J., Maki, T., Maksyotov, 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 Biogeochem. Cy., 18, GB1010, doi:10.1029/2003GB002111, 2004.

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., Kowalczyk, E., Maki, T., Maksyutov, S., Peylin, P., Prather, M., Pak, B. C., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: TransCom and CO_2 inversion intercomparison 1. Annual and mean control results and sensitivity to transport and prior flux information, Tellus 55B , 555-579, doi: 10.1034/j.1600-0889.2003.00049.x, 2003.

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., and et al.,: Towards robust regional estimates of CO_2 sources and sinks using atmospheric transport models, Nature 415 , 626-630, 2002.

Hammer, S., Glatzel-Mattheier, H., Müller, L., Sabasch, M., Schmidt, M., Schmitt, S., Schönherr, C., Vogel, F., Worthy, D. E., and Levin, I.: A gas chromatographic system for high-precision quasi-continuous atmospheric measurements of CO2, CH4, N2O, SF6, CO and H2, available at: "http://www.iup.uni-heidelberg.de/institut/forschung/groups/kk/GC_html" (last access: 25 January 2016), 2008.

Haszpra, L., Barcza, Z., Bakwin, P.S., Berger, B.W., Davis, K.J., Weidinger, T.: Measuring system for the long-term monitoring of biosphere/atmosphere exchange of carbon dioxide. J. Geophys Res, 106D, 3057-3069, DOI: 10.1029/2000JD900600, 2001.

Heimann, M. and Körner, S.: The global atmospheric tracer model TM3, Tech. Rep. 5, MPI BGC, Jena (Germany), online available at: http://www.bgc-jena.mpg.de/mpg/websiteBiogeochemie/Publikationen/Technical Reports/tech report5.pdf, 2003.

Helfter, C., Campbell, C., Dinsmore, K.-J., Drewer, J., Coyle, M., Anderson, M., Skiba, U., Nemitz, E., Billett, M.-F., and. Sutton, M.-A.: Drivers of long-term variability in CO2 net ecosystem exchange in a temperate peatland. Accepted for publication in Biogeosciences (March 2015).

Hilton, W-T., Davis, J-K., Keller, K., and Urban, M-N.: Improving terrestrial CO_2 flux diagnosis using spatial structure in land surface model residuals, Biogeosciences,

Houweling, S., Aben, I., Breon, F.-M., Chevallier, F., Deutscher, N., Engelen, R., Gerbig, C., Griffith, D., Hungershoefer, K., Macatangay, R., Marshall, J., Notholt, J., Peters, W., and Serrar, S.: The importance of transport model uncertainties for the estimation of CO_2 sources and sinks using satellite measurements, Atmos. Chem. Phys., 10, 9981-9992, doi:10.5194/acp-10-9981-2010, 2010.

Houweling, S., Breon, F.-M., Aben, I., Rödenbeck, C., Gloor, M., Heimann, M., and Ciais, P.: Inverse modeling of CO_2 sources and sinks using satellite data: A synthetic inter-comparison of measurement techniques and their performance as a function of space and time, Atmospheric Chemistry and Physics 4 , 523-538, 2004.

Hsieh, C. I., Katul, G. G., & Chi, T. W. (2000). An approximate analytical model for footprint estimation of scalar fluxes in thermally stratified atmospheric flows. Advances in Water Resources, 23, 765−772.

Jans, W.-W.-P., Jacobs, C.-M.-J., Kruijt, B., Elbers, J.-A.., Barendse, S., and Moors, E.-J.: Carbon exchange of a maize (Zea mays L.) crop: influence of phenology, Agriculture,

Ecosystems & Environment, 139, 316-324, 2010.

Elbers, J.-A., Jacobs, C.-M.-J., Kruijt, B., Jans, W.-W.-P., and Moors, E.-J.: Assessing the uncertainty of estimated annual totals of net ecosystem productivity: A practical approach applied to a mid latitude temperate pine forest, Agricultural and Forest Meteorology, Volume 151, Issue 12, 15 December 2011, Pages 1823-1830, ISSN 0168-1923, 10.1016/j.agrformet.2011.07.020.

Jarosz, N., Brunet, Y., Lamaud, E., Irvine, M., Bonnefond, J.-M. and Loustau, D.: Carbon dioxide and energy flux partitioning between the understorey and the overstorey of a maritime pine forest during a year with reduced soil availability, Agricultural and Forest Meteorology 148 1508-1523, 2008.

Jongen, M., Pereira, J., Saires., L.-M.-I., and Pio, C.-A.: The effects of drought and timing of precipitation on the inter-annual variation in ecosystem-atmosphere exchange in a Mediterranean grassland. Agricultural and Forest Meteorology.151, 595-606. http://dx.doi.org/10.1016/j.agrformet.2011.01.008, 2011.

Jung, M., Reichstein, M., and Bondeau, A.: Towards global empirical upscaling of FLUXNET eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, Biogeosciences, 6, 2001-2013, doi:10.5194/bg-6-2001-2009, 2009.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, Remote Sensing of Environment, 101(4), 534–553, doi:10.1016/j.rse.2006.01.020, 2006.

Kadygrov, N., Broquet, G., Chevallier, F., Rivier, L., Gerbig, C. and Ciais, P. (2015), On the potential of ICOS atmospheric CO 2 measurement network for the estimation of the biogenic CO_2 budget of Europe, *Atmospheric Chemistry and Physics Discussions* 15 , 14221-14273

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M. and Koffi, E.: Observing the continental-scale carbon balance assessment of and sampling complementarity and redundancy in a terrestrial and assimilation system by means of quantitative network design, Atmos. Chem. and Phys. 12 , 7867-7879 , doi: 10.5194/acp-12-7867-2012, 2012.

Kaminski, T. and Heimann, M.: A coarse grid three-dimensional global inverse model of the atmospheric transport 1. Adjoint model and Jacobian matrix, J. Geophys. Res. 104, D15, 18,535-18,553, doi: 10.1029/1999JD900147, 1999a.

Kaminski, T., Heimann, M. and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. Inversion of the transport of CO_2 in the 1980s, J. Geophys. Res.: Atmospheres (1984--2012) 104, D15, 18,555-18,581, doi: 10.1029/1999JD900146, 1999b.

Keenan, T., Garcia, R., Friend, A.-D., Zaehle, S., Gracia, C., Sabate, S.: Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO_2 and water fluxes through combined in situ measurements and ecosystem modeling, Biogeosciences, 6, 2285-2329, 2009.

Klumpp, K., Tallec, T., Guix, N., Soussana, J.-F.: Long-term impacts of agricultural practices and climatic variability on carbon storage in a permanent pasture, Global Change Biology, 17, 3534–3545, 2011.

Knohl, A., Schulze, E.-D., Kolle, O., Buchmann, N.: Large carbon uptake by an unmanaged 250-

year-old deciduous forest in Central Germany, Agricultural and Forest Meteorology 118, 151-167, 2003.

Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehrkorn, T., Daube, B. C., Kaplan, J. O., Houweling, S., and Wofsy, S. C. (2008). Emissions of ch4 and n2o over the united states and canada based on a receptor-oriented modeling framework and cobra-na atmospheric observations. Geophys. Res. Lett., 35(18):doi:10.1029/2008GL034031.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F.: World Map of the Köppen-Geiger climate classification updated, Meteorologische Zeitschrift, 15, 259-263, 2006.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO_2 inversions at the mesoscale using data driven prior uncertainties. Part1: Methodology and system evaluation, Atmos. Chem. Phys., acp-2016-577, submitted on 04 Jul 2016, 2016a.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, F. Th., Heimann, M.: Atmospheric CO2 inversions at the mesoscale using data driven prior uncertainties. Part2: the European terrestrial CO_2 fluxes, Atmos. Chem. Phys., acp-2016-578, submitted on 04 Jul 2016, 2016b.

Kountouris, P., Gerbig, C., Totsche, K. U., Dolman, A. J., Meesters, A. G. C. A., Broquet, G., Maignan, F., Gioli, B., Montagnani, L., Helfter, C.: An objective prior error quantification for regional atmospheric inverse applications, Biogeosciences, 12, 7403-7421, doi: 10.5194/bg-12-7403-2015, 2015.

Krinner, G., Viovz, N., Ogee, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, C.-I.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, Global Biogeochemical Cycles 19, doi: 10.1029/2003GB002199, 2005.

Kutsch, W.-L., Aubinet, M., Buchmann, N., et al.,: The net biome production of full crop rotations in Europe, Agriculture Ecosystems & Environment, 139, 336-345, 2010.

Lasslop, G., Reichstein, M., Kattge, J., and Papale, D.: Influences of observation errors in eddy flux data on inverse model parameter estimation, Biogeosciences Discussions 5, 2008.

Lauvaux, T., Schuh, A.-E., Bocquet, M., Wu, L., Richardson, S., Miles, N., and Davis, K.-J.: Network design for mesoscale inversions of CO_2 sources and sinks, Tellus B 64, doi: 10.3402/tellusb.v64i0.17980, 2012.

Lauvaux, T., Gioli, B., Sarrat, C., Rayner, P.-J., Ciais, P., Chevallier, F., Noilhan, J., Miglietta, F., Brunet, Y., Ceschia, E., Dolman, H., Elbers, J.-A., Gerbig, C., Hutjes, R., Jarosz, N., Legain, D., Uliasz, M.: Bridging the gap between atmospheric concentrations and local ecosystem measurements. Geophysical Research Letters, 36, Art. No. L19809, 2009.

Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K., Ciais, P., Denning, A. and Rayner, P.: Mesoscale inversion: first results from the CERES campaign with synthetic data, Atmos. Chem. and Phys., 8 , 3459-3471, doi:10.5194/acp-8-3459-2008, 2008.

Law, R. M., W. Peters, C. Rödenbeck, and TRANSCOM Contributors: TransCom model simulations of hourly atmospheric CO2: Experimental overview and diurnal cycle results for 2002, Global Biogeochem. Cycles, 22, GB3009, doi:10.1029/2007GB003050, 2008.

Lin, J. C., and C. Gerbig: Accounting for the effect of transport errors on tracer inversions, Geophys. Res. Lett., 32, L01802, doi:10.1029/2004GL021127, 2005.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J. Geophys. Res., 108, 4493, doi: 10.1029/2002JD003161, 2003

Lokupitiya, R. S., Zupanski, D., Denning, A. S., Kawa, S. R., Gurney, K. R. and Zupanski, M.: Estimation of global CO_2 fluxes at regional scale using the maximum likelihood ensemble filter, J. Geophys. Res. 113, D20110, doi: 10.1029/2007JD009679, 2008.

Longdoz, B., Gross, P., Granier, A.: Multiple quality tests for analysing CO_2 fluxes in a beech temperate forest. Biogeosciences. 2008, 5, 719–729.

Lopez, M., Schmidt, M., Ramonet, M., Bonne, J. L., Colomb, A., Kazan, V., Laj, P., and Pichon, J. M.: Three years of semicontinuous greenhouse gas measurements at the Puy de Dôme station (central France), Atmos. Meas. Tech., 8, 3941-3958, doi:10.5194/amt-8-3941-2015, 2015.

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

Marcolla, B., Cescatti, A., Manca, G., Zorer, R., Cavagna, M., Fiora, A., Gianelle, D., Rodeghiero, M., Sottocornola, M., Zampedri, R.: Climatic controls and ecosystem responses drive the inter-annual variability of the net ecosystem exchange of an alpine meadow, Agric. For. Meteorol., 151, 1233–1243, 2011.

Marcolla, B., Cescatti, A., Montagnani, L., Manca, G., Kerschbaumer, G., Minerbi, S.: Importance of advection in the atmospheric CO_2 exchanges of an alpine forest. Agric. For. Meteorol., 130, 193–206, 2005.

Marcolla, B., Pitacco, A., Cescatti, A.: Canopy architecture and turbulence structure in a coniferous forest, Boundary-Layer Meteorol., 39–59, 2003.

Matteucci, M., Gruening, C., Goded., B., I., Cescatti, A.: Soil and ecosystem carbon fluxes in a Mediterranean forest during and after drought. Agrochimica LVIII, 91–115, 2014.

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., Dolman, A. J.: Inverse carbon dioxide flux estimates for the Netherlands, J. Geophys. Res.-Atmos. 117, D20306, 1984-2012 , doi: 10.1029/2012jd017797, 2012.

Meijide, A., Manca, G., Goded, I., Magliulo, V., di Tommasi, P., Seufert, G., Cescatti, A.: Seasonal trends and environmental controls of methane emissions in a rice paddy field in Northern Italy, Biogeosciences 8, 3809–3821, 2011.

Michalak, A., Hirsch, A., Bruhwiller, L., Gurney, K. R., Peters, W., and Tans, P. P.: Maximum likelihood estimation of covariance parameters for Bayesian atmospheric trace gas surface flux inversions, J. Geophys. Res., 100, D24107, doi:10.1029/2005JD005970, 2005.

Michalak, A., M., Bruhwiler L. and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases, J. Geophys. Res. 109, D14109, doi: 10.1029/2003JD004422, 2004.

Mikaloff F., S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., Gloor, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Müller, S. A., and Sarmiento, J. L.: Inverse estimates of the oceanic sources and sinks of natural CO_2 and the implied oceanic transport, Glob. Biogeochem. Cy., 21, GB1010, doi:10.1029/2006GB002751, 2007.

Montagnani, L., Manca, G., Canepa, E., Georgieva, E., Acosta, M., Feigenwinter, C., Janous, D., Kerschbaumer, G., et al.,: A new mass conservation approach to the study of CO_2 advection in an alpine forest, Journal of Geophysical Research – Atmospheres, 114, D07306, DOI: 10.1029/2008JD010650, 2009.

Moors E.-J., Jacobs, C., Jans, W., Supit, I., Kutsch, W.-l, Bernhofer, C., Bezat, P., Buchmann, N., Carrara, A., Ceschia, E., Elbers, J., Eugster, W., Kruijt, B., Loubet, B., Magliulo, E., Moureaux, C., Olioso, A., Saunders, M., Soegaard, H.: Variability in carbon exchange of European croplands. Agriculture, Ecosystem & Environment, 139(3): 325-335, 2010

Moureaux, C., Debacq, A., Bodson, B., Heinesch, B., Aubinet, M.: Annual net ecosystem carbon exchange by a sugar beet crop. Agricultural and Forest Meteorology 139, 25-39, 2006.

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

Nagy, Z., Pintér, K., Czóbel, Sz., Balogh, J., Horváth, L., Fóti, Sz., Barcza, Z., Weidinger, T., Csintalan, Zs., Dinh, N., Q., Grosz, B., Tuba, Z.: The carbon budget of a semiarid grassland in a wet and a dry year in Hungary. AGR ECOSYST ENVIRON 121:(1-2)21-29, 2007.

Necki, J.M.,Chmura, L., Zimnoch, M., Rozanski, K.: Impact of emissions on atmospheric composition at Kasprowy Wierch based on results of carbon monoxide and carbon dioxide monitoring, Polish J. Environm. Studies, 22, 4, 1119-1127, 2013.

Olson, J.-S.: Global ecosystem framework-definitions, USGS EROS Data Cent. Intern. Rep. Sioux Falls SD, 37, 1994, 1994.

Papale, D. and Valentini, R., A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biology, 9: 525–535. doi: 10.1046/j.1365-2486.2003.00609.x, 2003

Patra, P. : TransCom model simulations of hourly atmospheric CO_2: Analysis of synoptic scale variations for the period 2002–2003, Global Biogeochem. Cycles, 22, GB4013, doi:10.1029/2007GB003081, 2008.

Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K., 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., Stroem, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, Rodo. X., Morgui, 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.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by

atmospheric observations, Glob. Change Biol., 16, 1317-1337, doi: 10.1111/j.1365-2486.2009.02078.x, 2010.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., B.Miller, J., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Wennberg, J. T. R. 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, 48, 18,925-18,930, doi: 10.1073pnas.0708986104, 2007.

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, 2013.

Peylin, P., Houweling, S., Krol, M. C., Karstens, U., Rödenbeck, C., Geels, C., Vermeulen, A., Badawy, B., Aulagnier, C., Pregger, T., Delage, F., Pieterse, G., Ciais, P., and Heimann, M.: Importance of fossil fuel emission uncertainties over Europe for CO_2 modeling: model intercomparison, Atmos. Chem. Phys., 11, 6607-6622, doi:10.5194/acp-11-6607-2011, 2011.

Peylin, P., Rayner, P., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P. and AEROCARB contributors: Daily CO_2 flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys. 5, 3173-3186, doi:10.5194/acp-5-3173-2005, 2005.

Piao, S., Fang, J., Ciais, P., Peylin, P., Huang, Y., Sitch, S., and Wang, T.: The carbon balance of terrestrial ecosystems in China, Nature 458, 1009-1013, doi: 10.1038/nature07944, 2009.

Pillai, D., Gerbig, C., Kretschmer, R., Beck, V., Karstens, U., Neininger, B., and Heimann, M.: Comparing Lagrangian and Eulerian models for CO_2 transport- a step towards Bayesian inverse modeling using WRF/STILT-VPRM, Atmospheric Chemistry and Physics 12 , 8979-8991 , doi: 10.5194/acp-12-8979-2012, 2012.

Pillai, D., Gerbig, C., Ahmadov, R., R¨odenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neininger, B., and Lavri´e, J. V.: High-resolution simulations of atmospheric CO_2 over complex terrain – representing the Ochsenkopf mountain tall tower. Atmos. Chem. Phys., 11(15):7445–7464, 2011.

Pita, G., Gielen, B., Zona, D., Rodrigues, A., Rambal, S., Janssens, I., Ceulemans, R.: Carbon and water vapor fluxes over four forests in two contrasting climatic zones, Agricultural and Forest Meteorology, 180, 211-224, 2013.

Popa, M. E., Gloor, M., Manning, A. C., Jordan, A., Schultz, U., Haensel, F., Seifert, T., and Heimann, M.: Measurements of greenhouse gases and related tracers at Bialystok tall tower station in Poland, Atmos. Meas. Tech., 3, 407-427, doi:10.5194/amt-3-407-2010, 2010.

Ramonet, M., Ciais, P., Aalto, T., Aulagnier, C., Chevallier, F., Cipriano, D., Conway, T. J., Haszpra, L., Kazan, V., Meinardt, F., Paris, J. D., Schmidt, M., Simmonds, P., Xueref-Remy, I. and Necki, J. N.: A recent build-up of atmospheric CO2 over Europe. Part 1: observed signals and possible explanations. Tellus B, 62: 1–13. doi: 10.1111/j.1600-0889.2009.00442.x, 2010

Prescher, A., K., Grünwald, T., Bernhofer, C.: Land use regulates carbon budgets in eastern Germany: From NEE to NBP, Agricultural and Forest Meteorology, 150, 1016-1025, 2010.

Rayner, P. J., Scholze, M., Knorr, W., Kaminski, T., Giering, R. and Widmann, H.: Two decades of terrestrial carbon fluxes from a carbon cycle data assimilation system (CCDAS), Glob. Biogeochem. Cy. 19 , GB2026 , doi: 10.1029/2004GB002254, 2005.

Rayner, P.J.: Optimizing CO_2 observing networks in the presence of model error: results from TransCom 3, Atmos. Chem. Phys., 4, 413-421, 2004.

Rayner (1996), Optimizing the CO_2 observing network for constraining sources and sinks, Tellus, 48B, 433-444, 1996.

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, Atmos. Chem. Phys., 14, 13739-13753, doi:10.5194/acp-14-13739-2014, 2014.

Rey, A., Belelli-Marchesini, L., Were, A., Serrano-Ortiz, P., Etiope, G., Papale, D., Domingo, F., and Pegoraro, E.: Wind as a main driver of the net ecosystem carbon balance of a semiarid Mediterranean steppe in the South East of Spain, Global Change Biology, 18, 539-554, 2012.

Richardson, A- D., Mahecha, M.-D., Falge, E., Kattge, J., Moffat, A.-M., Papale, D., Reichstein, M., Stauch, V.-J., Braswell, B.-H., Churkina, G., and et al.,: Statistical properties of random CO_2 flux measurement uncertainty inferred from model residuals, Agricultural and Forest Meteorology 148 , 38-50 , doi: 10.1016/j.agrformet.2007.09.001, 2008.

Richardson, A.-D., Hollinger, D.-Y., Burba, G.-G., Davis, K.-J., Flanagan, L.-B., Katul, G.-G., William M.-J., Ricciuto, D.-M., Stoy, P.-C., Suyker, A.-E., and et al.,: A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes, Agricultural and Forest Meteorology 136 , 1-18, 2006.

Rivier, L., Peylin, P., Ciais, P., Gloor, M., Roedenbeck, C., Geels, C., Karstens, U., Bousquet, P., Brandt, J. and Heimann, M.: European CO_2 fluxes from atmospheric inversions using regional and global transport models, Climatic Change, 103, 93-115, doi: 10.1007/s10584-010-9908-4, 2010.

Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., Reum, F., Keeling, R. F. and Heimann, M.: Interannual sea–air CO2 flux variability from an observation-driven ocean mixed-layer scheme, Biogeosciences, 11(17), 4599–4613, doi:10.5194/bg-11-4599-2014-supplement, 2014.

Rödenbeck, C., Gerbig, C., Trusilova, K. and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmos. Chem. and Phys. 9, 5331-5342, doi:10.5194/acp-9-5331-2009, 2009.

Rödenbeck, C.: Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, Technical Report 6, Max Planck Institute for Biogeochemistry, Jena, http://www.bgc-jena.mpg.de/mpg/websiteBiogeochemie/ Publikationen/Technical Reports/tech report6.pdf, 2005.

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

Rödenbeck C., Houwelling S., Gloor M., and Heinmann, M.: CO_2 flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport, Atmos. Chem. and Phys. 3, 1919-1964, doi: 10.5194/acp-3-1919-2003, 2003b.

Rodgers, C., D. Inverse methods for Atmosphere Sounding: Theory and Practice, World Sci., River Edge, N. J., 2000.

Sarmiento J. L. and Gruber N.: Sinks for anthropogenic carbon, Phys. Today, 55(8), 30–36, S-0031-9228-0208-010-9, 2002.

Schuh, A.-E., Denning, A.-S., Corbin, K.-D., Baker, I.-T., Uliasz, M., Parazoo, N., Andrews, A.-E., and Worthy, D.-E.-J.: A regional high-resolution carbon flux inversion of North America for 2004, Biogeosciences 7 , 1625-1644 , doi: 10.5194/bg-7-1625-2010, 2010.

Schuh, A. E., Denning, A. S., Uliasz, M. and Corbin, K. D.: Seeing the forest through the trees: Recovering large-scale carbon flux biases in the midst of small-scale variability, J. Geophys Res., 114, doi: 10.1029/2008JG000842, 2009.

Schulze, E. D., Ciais, P., Luyssaert, S., Schrumpf, M., Janssens, I. A., Thiruchittampalam, B., Theloke, J., Saurat, M., Bringezu, S., Lelieveld, J., Lohila, A., Rebmann, C., Jung, M., Bastviken, D., Abril, G., Grassi, G., Leip, A., Freibauer, A., Kutsch, W., Don, A., Nieschulze, J., Börner, A., Gash, J. H., and Dolman, A. J.: The European carbon balance. Part 4: integration of carbon and other trace-gas fluxes, Glob. Change Biology, 16, 1451-1469, doi: 10.1111/j.1365-2486.2010.02215.x, 2010.

Schwalm, C. R., Williams, W. A., Schaefer, K., Anderson, R., Arain, M. A.: A model-data intercomparison of CO2 exchange across North America: results from the North American carbon program site synthesis. J. Geophys. Res. 115, G00H05. DOI: 10.1029/2009JG001229, 2010.

Skiba, U., Jones, S.-K., Drewer, J., Helfter, C., Anderson, M., Dinsmore, K., McKenzie, R., Nemitz, E., and Sutton, M.-A.: Comparison of soil greenhouse gas fluxes from extensive and intensive grazing in a temperate maritime climate, Biogeosciences, 10, 1231-1241, 10.5194/bg-10-1231-2013, 2013.

Soussana, J.-F., Allard, V., Pilegaard, K., Ambus, P., Amman, C., Campbell, C., Ceschia, E., Clifton-Brown, J., Czöbel, Sz., Domingues, R., et al.,: Full accounting of the greenhouse gas (CO_2 , N_2O, CH_4) budget of nine European grassland sites, Agriculture, Ecosystems & Environment, 121, 121-134, 2007.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C. and Mukai, H.: The CO2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmos. Chem. Phys., 11(14), 6855–6870, doi:10.5194/acp-11-6855-2011, 2011.

Suni, T., Rinne, J., Reissel, A., Altimir, N., Keronen, P., Rannik, Ü., Dal Maso, M., Kulmala, M., and Vesala, T.: Long-term measurements of surface fluxes above a Scots pine forest in Hyytiälä, southern Finland, 1996-2001. Boreal Environment Res. 4, 287-301, 2003.

Tallec, T., Béziat, P., Jarosz, N., Rivalland V., and Ceschia E.: Crops water use efficiencies: comparison of stand, ecosystem and agronomical approaches, Agricultural and Forestry Meteorology, 168, 69 – 168, 69-81, 2013.

Taufarova, K., Havrankova, K., Dvorská, A., Pavelka, M., Urbaniak, M., Janous, D.: Forest

ecosystem as a source of CO_2 during growing season: relation to weather conditions. International Agrophysics, 28, 239-249, doi:10.2478/intag-2014-0013, 2014.

Thompson, R. L., Gerbig, C. and Rödenbeck, C.: A Bayesian inversion estimate of N_2O emissions for western Europe and the assessment of aggregation errors, Atmos. Chem. Phys., 11, 3443-3458, doi: 10.5194/acp-11-3443-2011, 2011.

Thompson, R. L., Manning, A. C., Gloor, E., Schultz, U., Seifert, T., Hänsel, F., Jordan, A., and Heimann, M.: In-situ measurements of oxygen, carbon monoxide and greenhouse gases from Ochsenkopf tall tower in Germany, Atmos. Meas. Tech., 2, 573-591, doi:10.5194/amt-2-573-2009, 2009.

Tolk, L. F., Dolman, A. J., Meesters, A. G. C. A. and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, Atmos. Chem. Phys., 11, 10349-10365, doi: 10.5194/acp-11-10349-2011, 2011.

Trusilova, K., Rödenbeck, C., Gerbig, C., and Heinmann, M.: Technical Note: A new coupled system for global to regional downscaling of CO_2 concentration estimation, Atmos. Chem. Phys. 10, 3205-3213, doi:10.5194/acp-10-3205-2010, 2010.

Trusilova, K., and Churkina, G.: The Terrestrial Ecosystem Model GBIOME-BGCv1, Max-Planck Institute for Biogeochemistry, Technical Report 14, http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-20689/tech_report14.pdf, 2008.

Vermeulen, A. T., Hensen, A., Popa, M. E., Bulk, W. C. M., Jongejan, P. A. C.: Greenhouse gas observations from Cabauw Tall Tower (1992-2010), Atmos. Meas. Tech., 4, 617-644, doi:10.5194/amt-4-617-2011, 2011.

Weedon, G.-P., Balsamo, G., Bellouin, N., Gomes, S., Best, M.-J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 2014.

Wei, S, Yi, C., Hendrey, G., Eaton, T., Rustic, G., Wang, S., Liu, H., Krakauer, N.-Y., Wang, W., Desai, A.-R., et al.,: Data-based perfect-deficit approach to understanding climate extremes and forest carbon assimilation capacity, Environmental Research Letters. 9(6):065002, 2014.

Wu, L., Bocquet, M., Lauvaux, T., Chevallier, F., Rayner, P., and Davis, K.: Optimal representation of source-sink fluxes for mesoscale carbon dioxide inversion with synthetic data, Journal of Geophysical Research: Atmospheres (1984--2012) 116, doi: 10.1029/2011JD016198, 2011.

Xiao, J. F., Zhuang, Q. L., Baldocchi, D. D., et al.: Estimation of net ecosystem carbon exchange for the conterminous United States by combining MODIS and AmeriFlux data, Agr. Forest Meteorol., 148(11), 1827–1847, 2008.

Zeeman, M., J., Hiller, R., Gilgen, A.-K., Michna, P., Plüss, P., Buchmann, N., Eugster, W.: Management and climate impacts on net CO_2 fluxes and carbon budgets of three grasslands along an elevational gradient in Switzerland, Agric. For. Meteorol., 150, 519-530, doi:

Zhang, S., Zheng, X., Chen, J. M., Chen, Z., Dan, B., Yi, X., Wang, L. and Wu, G.: A global carbon assimilation system using a modified ensemble Kalman filter, Geoscientific. Model Development., 8, 805-816, doi: 10.5194/gmd-8-805-2015, 2015.

Zupanski, D., Denning, A. S., Uliasz, M., Zupanski, M., Schuh, A. E., Rayner, P. J., Peters, W. and Corbin, K. D.: Carbon flux bias estimation employing Maximum Likelihood Ensemble Filter (MLEF), J. Geophys. Res., 112, D17107, doi: 10.1029/2006JD008371, 2007.

Acknowledgments

I could not start with anyone else than my “immediate” advisor Dr. Habil. Christoph Gerbig in my acknowledgements. I am deeply indebted for his continuous support of my Ph.D study and related research, for his patience, and immense knowledge. His friendliness, willingness and availability for various discussions helped me throughout my research and writing of this thesis. Sound advises, good ideas, his motivation and the constructive criticism inspired and helped me to bring this work to an end. I could not have imagined having a better advisor and mentor for my Ph.D. study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Kai-Uwe Totsche and Prof. Han Dolman for their insightful comments and encouragement during the regular Ph.D. progress meetings, as well as for their great assistance and comments while preparing the scientific papers.

I am much obliged to Prof. Dr. Martin Heimann for his valuable suggestions and support during crucial moments in my Ph.D. Further, I am very much thankful to Dr. Christian Rödenbeck and Dr. Ute Karstens for having their doors always open for fruitful discussions. I thank them for their amazing support and help which led to my full comprehension of the modeling framework I used, but also for their sound comments during the papers revision.

I would like to express also my deep appreciation to Dr. Thomas Koch for his assistance in programming and data retrievals. I thank him for the fruitful discussions and for the chance I had to learn so many diverse things. I am very thankful also to Dr. Antoon Meesters and Dr. Gregoire Broquet for all the data and help they provided for publishing concrete and scientifically sound papers.

All this work without the IT department would have never been achieved. I would like to thank the IT department at the Max Planck Institute in Jena and especially Dr. Peer-Joachim Koch for all of his assistance with respect to hardware/software problems, as well as his efforts to maintain the HPC cluster. I am also deeply thankful to Dr. Hendryk Bockelmann and the Deutsches KlimaRechenRentrum (DKRZ) supercomputing department. I thank them for allowing me to use the facilities, and their valuable help on solving software related problems.

My family’s generous and constant help and support gave me the strength to continue and overcome all the obstacles raised during this research. There are no words to thank my parents Kountouris Antonis and Kountouri Eleni for their support throughout my whole life.

Finally, I would like to express my gratitude to all of my friends that helped me to have a smooth life far away from home. Specifically, I am grateful to my friend Melanie Kirsch, who believed in me, while her altruistic support always filled me with the required courage to continue.

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und unter Verwendung der angegebenen Hilfsmittel, persönlichen Mitteilungen und Quellen angefertigt habe.

Jena, Kountouris, Panagiotis

CURRICULUM VITÆ

Personal information	
First name:	Panagiotis
Family name:	Kountouris
Date of birth:	29/01/1982
Address:	Leipziger Straße 18, Jena, Germany
Phone:	0049 176 79071136
Email:	pkount@bgc-jena.mpg.de
Education	
11/2011 - 08/2016	PhD researcher at Max Planck Institute for Biogeochemistry (Jena, Germany)
2007 - 2009	Master on Environmental Physics and Meteorology, Department of Physics, National Kapodistrian University of Athens
1999 - 2005	Undergraduate studies on Physics, Department of Physics, Aristotle University of Thessaloniki
Experience and Skills	
2011 - 2016	Researcher, Max Planck Institute for Biogeochemistry
Communication:	• Strong communication/presentation skills: presented data clearly both home or abroad • Adapting style and content to communicate results according to the audience level • Strong synthesizing and writing skills. Wrote scientific articles in international peer-reviewed journals. Wrote regular progression reports. Communicate results also to non expert
Interpersonal skills:	• Coached and collaborated with other scientists. Setting objectives, methodology, interpreting results • Collaborated and communicated at all professional levels, and people from diverse origins/cultures. Able to work as team player or independent • Experienced to analyze large amounts of data, check data quality, set strategy, analytical way of thinking to identify and to resolve problems

	• Managed several projects and collaborations. Used with deadlines, met targets on time.
2009	Research associate • Integrated management of industrial waste and sewage of Oinofuta Sximatarion Region (Supervisor: Prof. Maria Loizidou, National Metsobion University of Athens, chemical engineer department, in collaboration with the Greek Ministry of Environment Land Planning and Public Works.) • Design and Application of an Innovative Composting Unit for the Effective Treatment of Sludge and other Biodegradable Organic Waste in Morocco, MOROCOMP, LIFE-3rd Countries, 2006-2008, European Commission. (Supervisor: Prof. Maria Loizidou, National Metsobion University of Athens, chemical engineer department)

Jena, Kountouris, Panagiotis