



**HAL**  
open science

## Modélisation inverse des flux de CO<sub>2</sub> en Amazonie

Luis Molina Carpio

► **To cite this version:**

Luis Molina Carpio. Modélisation inverse des flux de CO<sub>2</sub> en Amazonie. Océan, Atmosphère. Université Paris Saclay (COMUE), 2017. Français. NNT : 2017SACLV040 . tel-01697602

**HAL Id: tel-01697602**

**<https://theses.hal.science/tel-01697602>**

Submitted on 31 Jan 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2017SACLV040

**MODÉLISATION INVERSE DES FLUX DE CO<sub>2</sub> EN  
AMAZONIE**

Thèse de doctorat  
de l'Université Paris-Saclay  
préparée à l'Université de Versailles Saint-Quentin-en-Yvelines

Ecole doctorale N° 129  
Sciences de l'Environnement en Ile-de-France (SEIF)  
Spécialité de doctorat: Météorologie, océanographie et physique de  
l'environnement

Thèse présentée et soutenue à Gif sur Yvette, le 24 octobre 2017, par

**M. LUIS MOLINA CARPIO**

Composition du Jury :

M. MATTHIEU ROY-BARMAN, LSCE, PRÉSIDENT  
M. EMANUEL GLOOR, UNIVERSITY OF LEEDS, RAPPORTEUR  
M. WOUTER PETERS, UNIVERSITY OF WAGENINGEN, RAPPORTEUR  
MME. KARLA LONGO, USRA/GESTAR, EXAMINATEUR  
M. BRUNO GUILLAUME, ARIA TECHNOLOGIES, EXAMINATEUR  
M. PHILIPPE CIAIS, LSCE, DIRECTEUR DE THÈSE  
M. GRÉGOIRE BROQUET, LSCE, CO-DIRECTEUR DE THÈSE



**Titre : Modélisation inverse des flux de CO<sub>2</sub> en Amazonie**

**Keywords :** Amazonie, flux CO<sub>2</sub> biogénique, inversion atmosphérique, échelle régionale

**Résumé :** Une meilleure connaissance des variations saisonnières et interannuelles du cycle du carbone en Amazonie est essentielle afin de comprendre le rôle de cet écosystème dans le changement climatique. La modélisation atmosphérique inverse est un outil puissant pour estimer ces variations, en exploitant l'information sur la distribution spatiale et temporelle des flux de CO<sub>2</sub> en surface contenue dans des observations de CO<sub>2</sub> atmosphériques. Néanmoins, la confiance en les estimations des flux en Amazonie obtenues à partir des systèmes d'inversion mondiale est faible du fait du manque d'observations dans cette région.

Dans ce contexte, j'ai d'abord analysé en détail les estimations de l'échange net de CO<sub>2</sub> entre la biosphère et l'atmosphère (NEE) générées par deux inversions mondiales pour la période 2002 – 2010. Ces deux inversions ont assimilé des données provenant du réseau mondial d'observation du CO<sub>2</sub> atmosphérique hors de l'Amérique du Sud, et une d'elles a assimilé des observations de quatre stations de surface en Amazonie, qui n'ont jamais été exploitées dans les études d'inversion précédentes. J'ai montré que dans une inversion mondiale les observations de stations loin d'Amazonie et les observations locales contrôlaient la NEE. Pourtant, les résultats ont révélé des structures à très grande échelle peu réalistes. L'analyse a confirmé le manque de stations en Amazonie pour fournir des estimations fiables, et les limites des systèmes d'inversion mondiale avec des modèles à très basse résolution.

J'ai donc ensuite évalué l'apport de l'utilisation du modèle atmosphérique régional BRAMS, par rapport à celle du système mondial de prévision météorologique ECMWF, pour le forçage météorologique du modèle de transport atmosphérique CHIMERE simulant le CO<sub>2</sub> en Amérique du Sud à haute résolution (~35 km). J'ai simulé le CO<sub>2</sub> avec les deux modèles de transport—CHIMERE-BRAMS et CHIMERE-ECMWF. J'ai évalué ces simulations avec des profils verticaux de mesures aéroportées, en analysant les mesures individuelles et les gradients horizontaux de CO<sub>2</sub> calculés entre paires de stations dans le sens du vent, à différentes altitudes ou intégrés sur la verticale. Les deux modèles de transport ont simulé les observations de CO<sub>2</sub> avec une performance similaire, mais j'ai trouvé une importante incertitude sur les modèles de transport. Les mesures individuelles et les gradients horizontaux ont été surtout sensibles à la NEE, mais aussi, pendant la saison sèche, aux émissions des feux de biomasse (E<sub>FIRE</sub>). J'ai trouvé que l'assimilation des gradients horizontaux était plus approprié pour les inversions que celle des mesures individuelles, étant donné que les premiers ont été moins sensibles au signal associé aux flux hors de l'Amérique du Sud et à l'incertitude sur le modèle de transport en altitude.

Finalement, j'ai développé deux systèmes d'inversion régionale pour l'Amérique du Sud tropicale avec les deux modèles de transport, et j'ai lancé des inversions avec quatre types de vecteurs d'observation: de mesures individuelles et gradients horizontaux sur cinq niveaux verticaux, à la surface, ou de gradients



horizontaux intégrés sur la verticale. J'ai trouvé une forte dépendance des estimations des bilans régionaux et sub-régionaux de NEE et  $E_{\text{FIRE}}$  au modèle de transport, ainsi qu'au vecteur d'observation. Les inversions assimilant des gradients horizontaux ont séparé mieux les signaux de NEE et  $E_{\text{FIRE}}$ . Cependant, les grandes incertitudes sur les flux inversés ont réduit la confiance en ces estimations. Par conséquent, si mon étude n'a pas amélioré la connaissance des variations saisonnières et interannuelles de la NEE en Amazonie, elle a montré les besoins d'amélioration de la modélisation du transport dans la région et de la stratégie de modélisation inverse, du moins à travers une définition du vecteur d'observation appropriée qui prenne en compte les caractéristiques des données disponibles, et les limitations des modèles de transport actuels.



**Title : Inverse modeling of CO<sub>2</sub> fluxes in Amazonia**

**Keywords :** Amazonia, net ecosystem exchange, atmospheric inversion, regional scale

**Abstract :** A better knowledge of the seasonal and inter-annual variations of the Amazon carbon cycle is critical to understand the influence of this terrestrial ecosystem on climate change. Atmospheric inverse modeling is a powerful tool to estimate these variations by extracting the information on the spatio-temporal patterns of surface CO<sub>2</sub> fluxes contained in observations of atmospheric CO<sub>2</sub>. However, the confidence in the Amazon flux estimates obtained from global inversion frameworks is low, given the scarcity of observations in this region.

In this context, I started by analyzing in detail the Amazon net ecosystem exchange (NEE) inferred with two global inversions over the period 2002 – 2010. Both inversions assimilated data from the global observation network outside Amazonia, and one of them also assimilated data from four stations in Amazonia that had not been used in previous inversion efforts. I demonstrated that in a global inversion the observations from sites distant from Amazonia, as well as local observations, controlled the NEE inferred through the inversion. The inferred fluxes revealed large-scale structures likely not consistent with the actual NEE in Amazonia. This analysis confirmed the lack of observation sites in Amazonia to provide reliable flux estimates, and exposed the limitations of global frameworks, using low-resolution models to quantify regional fluxes. This limitations justified developing a regional approach.

Then I evaluated the benefit of the regional atmospheric model BRAMS, relative to the global forecast system ECMWF, when both models provided the meteorological fields to drive the atmospheric transport model CHIMERE to simulate CO<sub>2</sub> transport in tropical South America at high resolution (~35 km). I simulated the CO<sub>2</sub> distribution with both transport models—CHIMERE-BRAMS and CHIMERE-ECMWF. I evaluated the model simulations with aircraft measurements in vertical profiles, analyzing the concentrations associated to the individual measurements, but also with horizontal gradients along wind direction between pairs of sites at different altitudes, or vertically integrated. Both transport models simulated the CO<sub>2</sub> observations with similar performance, but I found a strong impact of the uncertainty in the transport models. Both individual measurements and horizontal gradients were most sensitive to NEE, but also to biomass burning CO<sub>2</sub> emissions ( $E_{\text{FIRE}}$ ) in the dry season. I found that horizontal gradients were more suitable for inversions than individual measurements, since the former were less sensitive fluxes outside South America and further decreased the impact of the transport model uncertainty in altitude.

Finally, I developed two analytical regional inversion systems for tropical South America, driven with CHIMERE-BRAMS and CHIMERE-ECMWF, and made inversions with four observation vectors: individual concentration measurements and horizontal gradients at five vertical levels, close to the surface, or horizontal gradients vertically integrated. I found a strong dependency of the inverted regional and sub-regional NEE and  $E_{\text{FIRE}}$  emissions budgets on both the transport model and the observation vector. Inversions with gradients yielded a better separation of NEE and  $E_{\text{FIRE}}$  signals. However, the large uncertainties in the inverted fluxes, did not yield high confidence in the estimates. Therefore, even though



my study did not improve the knowledge of seasonal and year-to-year variations of the NEE in Amazonia, it demonstrated need the of further efforts to improve transport modeling in the region and the inverse modeling strategy, at least through an appropriate definition of the observation vector that accounts for the characteristics of the available data, and the limitations of the current transport models.





---

# Contents

- 1 CO<sub>2</sub> exchanges by land ecosystem in Amazonia inferred from atmospheric inversion** **1**
  - 1.1 Principles of the inverse problem . . . . . 2
    - 1.1.1 Components of the inversion . . . . . 4
    - 1.1.2 Bayesian approach to the inverse problem . . . . . 10
    - 1.1.3 Solution of the inverse problem . . . . . 12
  - 1.2 Estimates of CO<sub>2</sub> fluxes in Amazonia from atmospheric inversion . . . . . 13
  - 1.3 Toward robust inverse modeling estimates of CO<sub>2</sub> fluxes in Amazonia . . . . . 17
  - 1.4 Structure of this work . . . . . 19
  
- 2 On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia** **21**
  - 2.1 Objective of the study . . . . . 21
  - 2.2 Main results and implications . . . . . 22
    - 2.2.1 Simulated vs. observed concentrations . . . . . 22
    - 2.2.2 Impact on surface biogenic CO<sub>2</sub> fluxes . . . . . 22
  - 2.3 Conclusions . . . . . 24
  
- 3 Regional atmospheric modeling of CO<sub>2</sub> transport in Amazonia** **25**
  - 3.1 Materials and methods . . . . . 27
    - 3.1.1 Transport configuration . . . . . 27
    - 3.1.2 Evaluation of simulated meteorological fields . . . . . 31
    - 3.1.3 Spatial and temporal CO<sub>2</sub> distribution and comparison to the regular vertical profiles from GA2014 . . . . . 34
  - 3.2 Results . . . . . 36
    - 3.2.1 Simulated vs. observed meteorology . . . . . 36
    - 3.2.2 Spatial and temporal CO<sub>2</sub> distribution . . . . . 41
  - 3.3 Discussion . . . . . 63
  - 3.4 Conclusions . . . . . 66
  
- 4 Regional atmospheric inversion of CO<sub>2</sub> in Amazonia** **69**

4.1	Inverse modeling setup . . . . .	70
4.1.1	Control vector . . . . .	71
4.1.2	Flux products used in the inverse modeling framework and definition of the control regions . . . . .	71
4.1.3	Prior uncertainty covariance structure . . . . .	78
4.1.4	Observation operator . . . . .	78
4.1.5	Observation vector and error covariance structure . . . . .	79
4.2	Results: Fit to observed CO <sub>2</sub> data . . . . .	81
4.3	Results: optimized fluxes . . . . .	94
4.4	Discussion . . . . .	102
4.5	Conclusions . . . . .	108
<b>5</b>	<b>Conclusions and perspectives of future research</b>	<b>111</b>
5.1	Conclusions . . . . .	112
5.2	Perspectives . . . . .	115
<b>A</b>	<b>Publication: On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia</b>	<b>117</b>
<b>B</b>	<b>Supplementary figures</b>	<b>135</b>
	<b>Bibliography</b>	<b>139</b>

---

# CO<sub>2</sub> exchanges by land ecosystem in Amazonia inferred from atmospheric inversion

Atmospheric carbon dioxide (CO<sub>2</sub>), along with methane (CH<sub>4</sub>) and nitrous oxide (N<sub>2</sub>O), exerts strong influence on the radiative properties of the atmosphere, which influences the global energy budget, and in turn, the Earth's climatic system. The abundance of CO<sub>2</sub> in the atmosphere has in the Industrial Era (since 1750; [IPCC, 2013](#)), mainly due to fossil fuel burning (coal, gas, oil and gas flaring), cement production, and land use change (mainly deforestation). Between 1750 and 2011, these human activities have released  $\sim 555 \pm 85$  PgC of anthropogenic carbon. According to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, it is virtually certain that this anthropogenic forcing has warmed the global climate system.

Less than half of the anthropogenic carbon emitted since 1750 has remained in the atmosphere; the rest has been absorbed by the ocean and terrestrial ecosystems (the carbon sinks). In the terrestrial land sink, Amazon forests are of particular interest because of the vast extension of this ecosystem and the large amount of carbon stored mainly in biomass. Rising CO<sub>2</sub> concentration fosters photosynthetic activity through the fertilization effect ([Amthor, 1995](#); [Cao and Woodward, 1998](#)) and modifies other physiological properties of canopy, with effects on the terrestrial water cycle (e.g. [Fatichi et al., 2016](#)). In combination with the high productivity rates and long carbon residence times in biomass in Amazon forests, this ecosystem could contribute to moderate CO<sub>2</sub> in the atmosphere, and mitigate global warming. On the other hand, if the carbon stored in Amazon forests and soils is released, for instance, as a consequence of biome shifts, due to hotter and drier climate (as projected by some climate change scenarios) future warming could be exacerbated.

The land carbon cycle is very sensitive to climate changes, e.g. precipitation, temperature. Thus, many research try to better estimate and understand the temporal and spatial variations of the Amazon land balance, in order to project the future of this ecosystem, and consequently of the global climate system. Such efforts range from local-scale ground-based studies, e.g. eddy-covariance measurements ([Restrepo-Coupe et al., 2013](#); [Saleska et al., 2003](#)) and forest inventories ([Phillips et al., 2009](#)), to larger scale e.g. through satellite observations. ([Gloor et al., 2012](#)) provided a review of some of these methods and their results. Remote-sensing observations of canopy greenness ([Huete et al., 2006](#); [Saleska et al., 2007](#); [Samanta et al., 2010](#)), canopy structure ([Saatchi et al., 2012](#)), fluorescence ([Lee et al., 2013](#); [Parazoo et al., 2013](#)), land use change ([Song et al., 2015](#)) fire

activity (Giglio et al., 2013; Kaiser et al., 2012; Wiedinmyer et al., 2011) have been used to constrain the seasonal and interannual variations of the Amazon land sink, and its response to both natural and anthropogenic disturbance. Nevertheless, this region encompasses a variety of climatic sub-regions and high spatial heterogeneity in species composition and physiognomy (Tuomisto et al., 1995; Xu et al., 2015), strong small-scale geomorphologic and edaphic differences. This complicates the integration of different observations across scales to provide a basin-scale view of the Amazon carbon balance, and makes carbon cycle modeling challenging in this region.

Another constraint of the Amazon carbon balance at very large spatial scales is given by observations of the spatial and temporal patterns of atmospheric CO<sub>2</sub> distribution, which reflects the spatio-temporal distribution of CO<sub>2</sub> sources and sinks integrated by atmospheric transport (Enting, 2002). This concept underlies the principle of atmospheric inverse modeling techniques, which are used to deduce CO<sub>2</sub> sources and sinks from concentration measurements. Inverse modeling has been used since the 1980s to deduce surface CO<sub>2</sub> fluxes latitudinal to continental scales (Bousquet et al., 2000; Fan et al., 1998), sub-continental scales (Peters et al., 2005; Rödenbeck et al., 2003; Rödenbeck et al., 2003), and more recently at regional scale inversion activities have emerged during the last 5 – 8 years (Broquet et al., 2011; Lauvaux et al., 2008; Rödenbeck et al., 2009).

My PhD thesis, presented in this dissertation, has been dedicated to the study of the Amazon CO<sub>2</sub> balance using atmospheric inverse modeling. This chapter introduces the rationale and specificities of this thesis. It gives the principles of the atmospheric inversion method, and the ingredients of an inversion system to solve for CO<sub>2</sub> sources and sinks (Section 1.1.1) and the formalism to express the inverse problem (Section 1.1.2), followed by an overview of the information that state-of-the-art inverse modeling systems yield about the Amazon carbon balance (Sections 1.2 and 1.3). Finally, it presents the objectives of my research and guides the reader through the rest of the thesis works developed over nearly four years, and synthesized in the subsequent chapters (Section 1.4).

## 1.1 Principles of the inverse problem

An inverse problem consists in finding the best estimates of parameters of a system based on observations of a measurable manifestation of that system. This is the principle of the estimation of CO<sub>2</sub> sources and sinks using measurements of atmospheric CO<sub>2</sub> concentration.

Let us consider a perfect atmospheric transport model  $h(x)$  that takes the true surface CO<sub>2</sub> flux  $x_t$  as parameter and predicts the true atmospheric CO<sub>2</sub> concentration  $y_t$  at a given point. One can write this as:

$$y_t = h(x_t) \tag{1.1}$$

$h(x)$  is a forward function (Rodgers, 2000) that encapsulates the relationship between fluxes and concentrations. However, the physics behind atmospheric transport is so complex so that one has to approximate the real processes by some forward model  $H(x)$ , which is a numerical model of atmospheric transport.

In general, true fluxes and concentrations are unknown. We may have, however, a prior estimate of the flux  $x_b$  bearing an uncertainty  $\varepsilon_b$ . With an imperfect model (i.e., a model with errors) one could only simulate a concentration  $y_b$  with an uncertainty  $\xi_b$  due to the errors both in  $x_b$  the initial parameter and the model. If one measures the true

concentration, this is an observation  $y_o$ , which bears an  $\epsilon_o$  error due instrumental errors. The true values should be within the uncertainty of the estimate  $y_o$ . This is illustrated in Figure 1.1a.

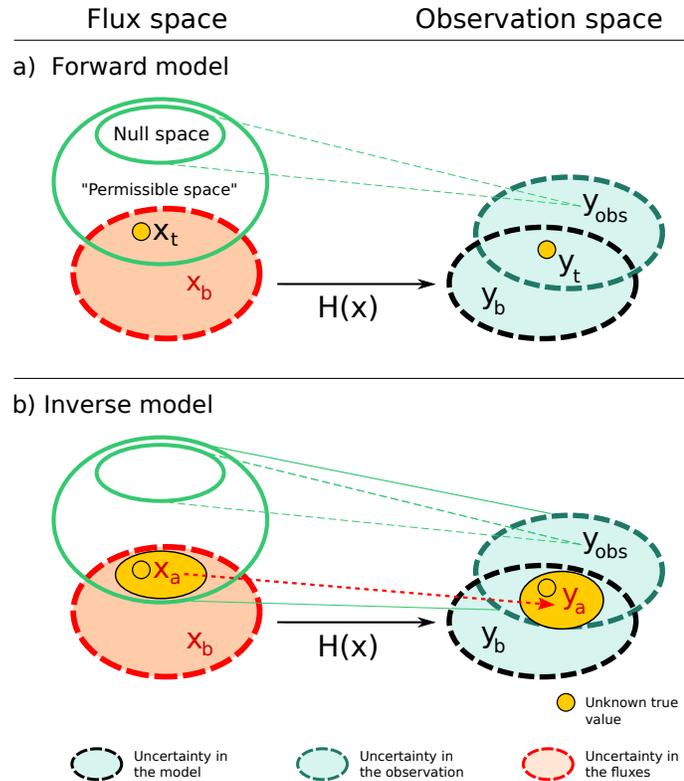


Figure 1.1: Inversion principle. Adapted from [Ciais et al. \(2010\)](#)

Adjusting the parameter  $x$  from the prior value  $x_b$ , so as to minimize the distance between the simulated and observed concentrations within their errors is an inversion, as illustrated in Figure 1.1b.  $x_a$  is the updated, optimized, or *posterior* flux and the corresponding simulated concentration is  $y_a$ . An important aspect of the inversion is that it brings an uncertainty reduction on both the updated fluxes and the updated concentration in comparison to their prior values, as a result of the information contained in the observations.

In the real world, fluxes are continuous and rather complex functions of space and time. This means that we are dealing with infinite very large number of unknown variables. To simplify the problem, a representation of the surface fluxes in terms of a finite and numerically tractable number of parameters. For instance, spatially, fluxes can be aggregated at small scales into a discrete number of regions, or represented as a gridded surface. Temporally, fluxes can also be aggregated over fixed time periods. But even after aggregation we usually end up with insufficient information because the number of available in-situ  $\text{CO}_2$  measurements is usually smaller than the number of unknown fluxes to determine.

Estimating  $\text{CO}_2$  sources and sinks using observations is thus an under-constrained problem. Mathematically, this implies that there is an infinite number of estimates of the parameters allowing to match the observations, as shown in Figure 1.1. When observations are scarce, not sensitive enough to the sought fluxes, or uncertain, the null space becomes

very large. In this case, optimizing the fluxes  $x$  to match the observations means to choose the parameters from a large null space that are consistent with the preexisting information  $x_b$ , considering uncertainties in the observations and  $x_b$ . From Figure 1.1 it follows that proper treatment of the uncertainties involved is essential to solve for the inverse problem. Therefore, a formalism is required to express the uncertainty in the input information and in the optimized parameters, and make sure that the latter has been reduced as much as possible. This formalism is generally provided by the Bayes' theorem, discussed in Section 1.1.2.

### 1.1.1 Components of the inversion

Figure 1.1 introduced the typical components of an inversion that estimates the optimal spatio-temporal distribution of the net surface CO<sub>2</sub> fluxes that best fits a set of observations of atmospheric CO<sub>2</sub> (taking into account the uncertainties in the prior fluxes and in the observations) using an atmospheric transport model. In the following, I discuss the components of an inversion system aimed at inferring land natural fluxes at large scales (i.e. continental to regional scale).

#### Prior estimate of CO<sub>2</sub> fluxes

The net CO<sub>2</sub> exchange, and consequently the atmospheric CO<sub>2</sub> content, is the result of both natural and anthropogenic processes that release and take up CO<sub>2</sub>. The contribution of these processes must be included in the prior fluxes to provide for the best possible initial estimate of net CO<sub>2</sub> exchange.

CO<sub>2</sub> released through combustion of carbon stored in the solid Earth by human activities represents a major perturbation of the carbon cycle (IPCC, 2013), and need to be included as prior information. Inventories of fossil fuel CO<sub>2</sub> emissions are based on data of fuel production or fuel consumption. Although global total anthropogenic CO<sub>2</sub> emissions from fuel combustion and cement production are known within 10% (Andres et al., 2012), uncertainty on the total emission at individual country level can reach more than 50% (Andres et al., 2012). Global maps of fossil fuel CO<sub>2</sub> emissions (e.g. EDGAR, Olivier et al., 2005); CDIAC (Andres et al., 1996); PK-CO2 (Wang et al., 2013) are based on national inventories, usually reported as tabulated data. These tables are usually translated into maps using either a proxy variable (e.g. population density, Andres et al., 1996) or variables (e.g. road density, energy consumption data by sector, urban-rural population density; Olivier et al., 2005), with better resolution than that of the original inventories, or using a process-based approach in which process data play an equivalent role than the proxy variable. Figure 1.2 shows an example of the global distribution of anthropogenic CO<sub>2</sub> emissions (Emissions Database for Global Atmospheric Research version 4.2; EU-JRC/PBL, 2013).

Biomass burning emissions are also an important source of CO<sub>2</sub>, as well as aerosols and many other chemical species. The usual approach to calculate these emissions is to estimate initially the amount of biomass actually burned, using information on the aboveground biomass density, the fraction of fuel load consumed by combustion (combustion factor), and the surface area burned. Then emission factors (ratio of the mass of the chemical species emitted to the total dry mass of fuel burned) are applied to estimate the emission of a compound. For CO<sub>2</sub>, the emission factor is close to one. It must be noticed that to the important exception of deforestation and peat fires, biomass burning emissions of

grasslands are thought to be compensated by ecosystem recovery, or regrowth, on time scales of a few years.

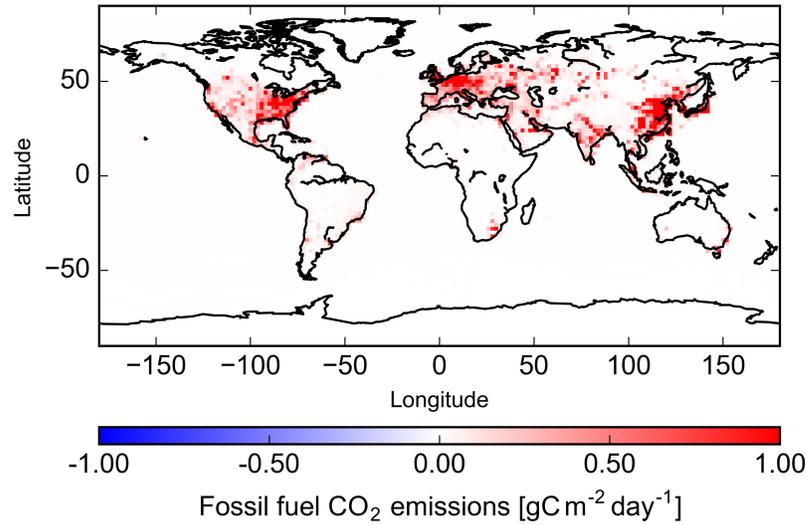


Figure 1.2: Annual fossil fuel CO<sub>2</sub> emissions for 2010 estimated from statistics on energy consumption. Emissions are estimated for each country, and then spatially distributed using information on location of energy and manufacturing facilities, road networks, shipping routes, human and animal population density and agricultural land use. Source: EDGAR version 4.2 FastTrack database, <http://edgar.jrc.ec.europa.eu>.

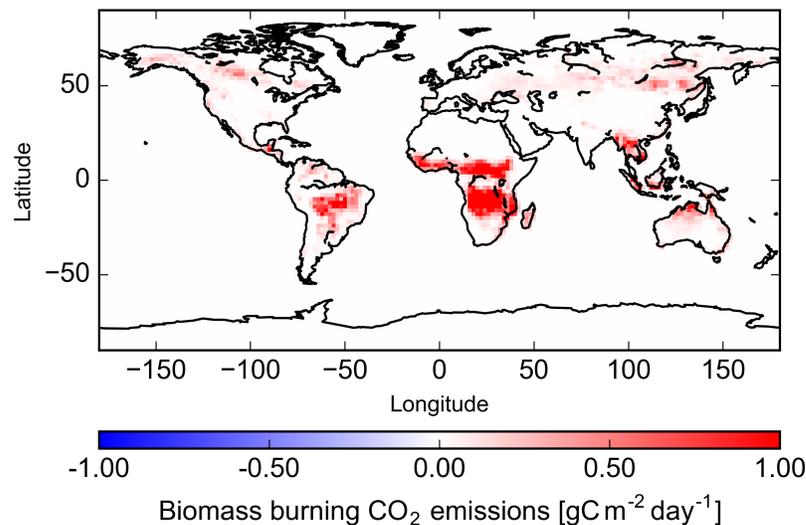


Figure 1.3: Distribution of mean biomass burning CO<sub>2</sub> emissions averaged over 1997 – 2014, based on a biogeochemical model and satellite-derived estimates of area burned, fire activity, and plant productivity (after Giglio et al., 2013). Source: <http://www.globalfiredata.org/figures.html>.

An example of the global distribution of CO<sub>2</sub> emissions from biomass burning is shown in Figure 1.3 (Global Fire Emission Database version 4s; Giglio et al., 2013).

CO<sub>2</sub> exchanges between the ocean and the atmosphere depend on the difference in the partial pressure of CO<sub>2</sub> ( $p\Delta\text{CO}_2$ ) between the air and the water and on gas exchange transfer velocities. Millions of measurements of surface water pCO<sub>2</sub> were collected across the oceans, and compiled in databases (e.g. Bakker et al., 2014)(Takahashi et al., 2016). These data sets, together with empirical formulations of air-sea gas transfer velocities, have been used to upscale spatio-temporal patterns of the air-sea CO<sub>2</sub> exchanges through a of interpolation methods (Landschützer et al., 2015; Rödenbeck et al., 2015). In data-scarce regions like the southern oceans (see Figure 1B in Takahashi et al., 2009), interpolated fields are the only information available about spatio-temporal distribution of ocean fluxes, and different products show significant differences. Figure 1.4 depicts a map of CO<sub>2</sub> air-sea exchange from Takahashi et al. (2009).

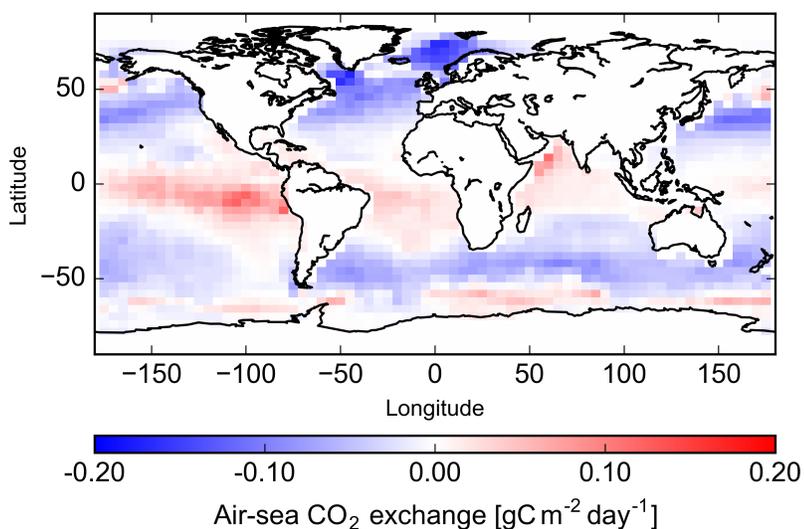


Figure 1.4: Climatological mean annual ocean CO<sub>2</sub> flux for the reference year 2000, for non-ENSO conditions (after Takahashi et al., 2009).

The net CO<sub>2</sub> exchange between the terrestrial biosphere and the atmosphere results from the CO<sub>2</sub> uptake by photosynthesis, and CO<sub>2</sub> release through the respiration of plants, heterotrophic organisms and fires. Furthermore, in some ecosystems also photo-oxidation releases CO<sub>2</sub> (Rutledge et al., 2010). In my PhD thesis, I have considered that carbon is exchanged as CO<sub>2</sub> with the atmosphere from photosynthesis, fires and plant and soil respiration, ignoring CO<sub>2</sub> exchanged between the land and the atmosphere from lateral fluxes generated from harvested wood and crop products, and from river transport (Ciais et al., 2008). Depending on the spatial scale involved, natural disturbances, such as fires, squall events, and pests can damage an ecosystem, releasing CO<sub>2</sub> immediately and with some delay. Anthropogenic disturbances such as forest clearing for agriculture also release important amounts of CO<sub>2</sub> to the atmosphere (van der Werf et al., 2009). Figure 1.5 shows an example of prior information on the net CO<sub>2</sub> flux between the land and the atmosphere from the dynamic global vegetation model (DGVM) ORCHIDEE (Krinner et al., 2005). DGVMs are process models that represent dynamically the energy, water and carbon exchanges between the terrestrial ecosystem and the atmosphere, and which are capable of

simulating vegetation changes in response to climatic changes, by simulating. Such DGVM estimates bear large uncertainties reflecting the incomplete knowledge of the underlying mechanisms that determine the net CO<sub>2</sub> exchange. Uncertainties stem from the extreme heterogeneity of the land surface (Ciais et al., 2010), and the complications to extrapolate the knowledge of ecosystem functioning, acquired from laboratory experiments or ground-based, local-scale observations of these fluxes to large-scale models (Bonan, 2008; Medlyn et al., 2015).

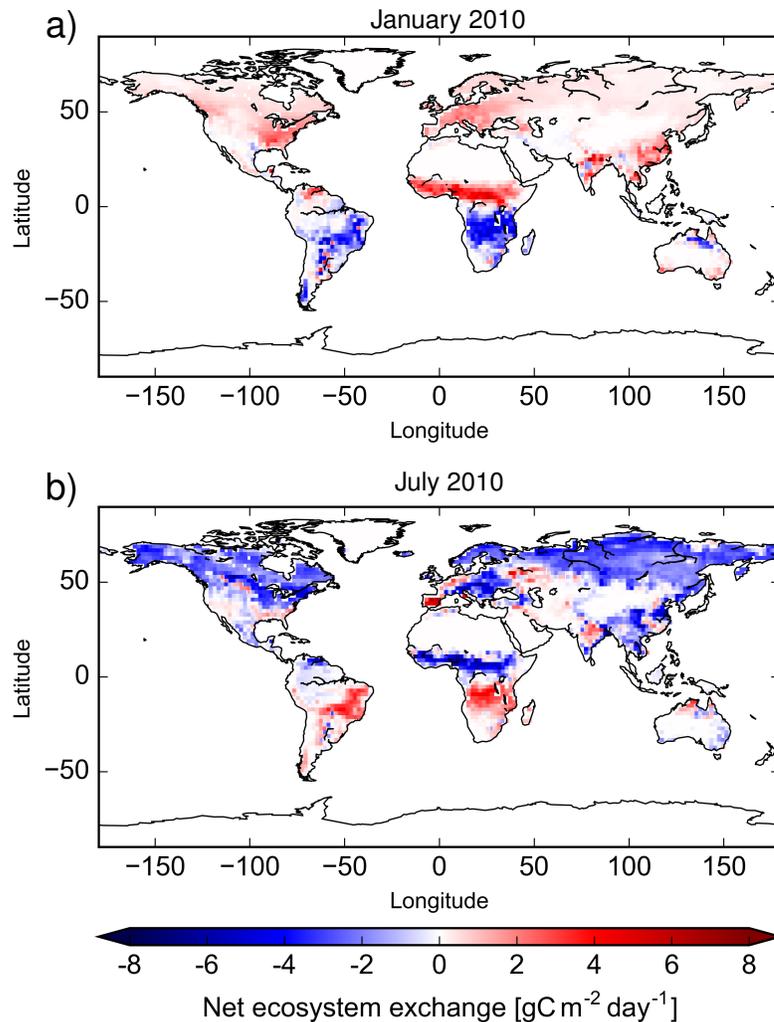


Figure 1.5: Monthly net ecosystem exchange for January (a) and July (b) 2010 simulated with the model ORCHIDEE. Fluxes are reported from the atmospheric perspective, such that positive values indicate emission to the atmosphere and negative values indicate up-take by the ecosystem.

### Atmospheric CO<sub>2</sub> observations

Measurements of atmospheric CO<sub>2</sub> date back to the first two stations in the South Pole and Mauna Loa, Hawaii in 1957 and 1958, respectively (Bolin and Keeling, 1963; Keeling,

# 1. CO<sub>2</sub> EXCHANGES BY LAND ECOSYSTEM IN AMAZONIA INFERRED FROM ATMOSPHERIC INVERSION

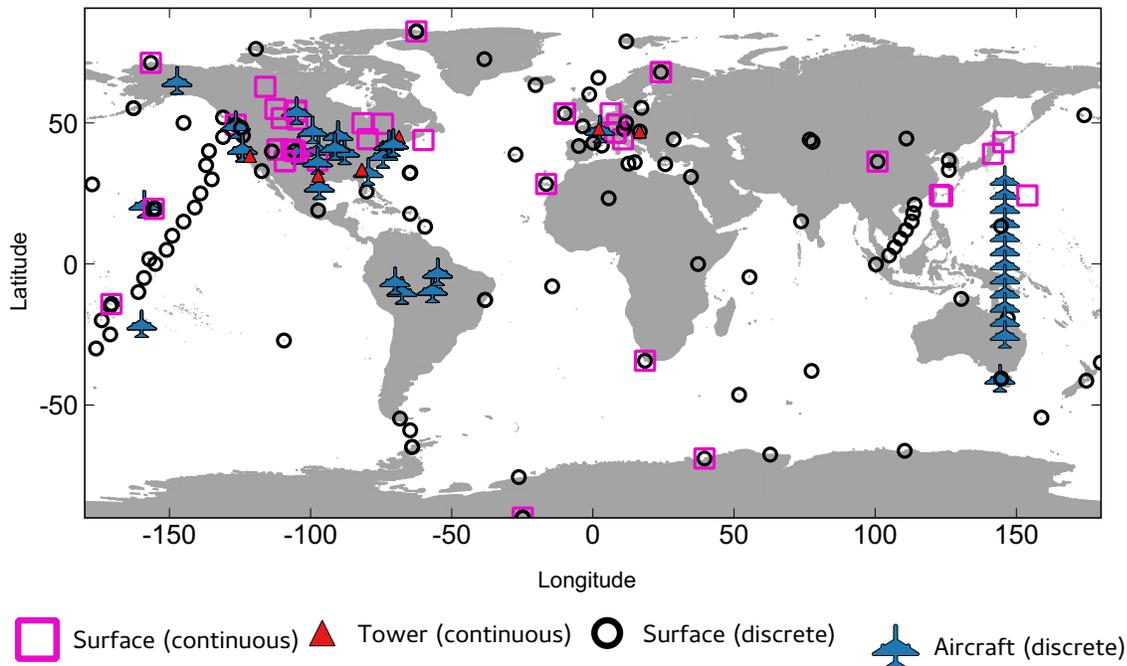


Figure 1.6: Sampling locations for measurement records used to derive GLOBALVIEW-CO<sub>2</sub>. Source: [http://www.esrl.noaa.gov/gmd/ccgg/globalview/co2/co2\\_observations.html](http://www.esrl.noaa.gov/gmd/ccgg/globalview/co2/co2_observations.html).

1960). Currently, atmospheric in-situ CO<sub>2</sub> observations are carried out at about 100 sites around the globe contributed by different laboratories (Ciais et al., 2010). Continuous measurements are made at some 30 sites, at surface stations and tall towers. Discrete flask air samples taken at surface stations or using airborne platforms add approximately 100 sites to the global network. Figure 1.6 illustrates the cooperative effort of multiple institutions to the observation of atmospheric CO<sub>2</sub> (GLOBALVIEW-CO<sub>2</sub>, 2013).

Figure 1.6 shows that large geographic land and ocean areas are under-sampled by the current in-situ network. The figure illustrates that most observation sites are located over the ocean because CO<sub>2</sub> concentrations in these areas are subject to less variability from smoother ocean fluxes given atmospheric transport within the boundary layer less variable than overland. This makes marine stations suitable to capture the signature of large-scale fluxes, but they cannot provide detailed constraints on land CO<sub>2</sub> surface fluxes. On land, North America and Europe have the most dense observation networks, but large areas of Asia, Africa, and South America remain mostly unconstrained in atmospheric inversions.

Satellite-based measurements of CO<sub>2</sub> offer a means to overcome the limited spatial coverage of the current observing. CO<sub>2</sub> retrievals have been obtained from radiances measured with instruments not originally designed for this purpose, e.g. the Television Infrared Observation Satellite Operational Vertical Sounder (TOVS), the Tropospheric Emission Spectrometer (TES) (Kulawik et al., 2010), Interferometric Atmospheric Sounding Instrument (IASI) (Crevoisier et al., 2009), and the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) (Buchwitz et al., 2007). The first CO<sub>2</sub>- and CH<sub>4</sub>-dedicated platform, the Greenhouse Gases Observing Satellite (GOSAT) (Kuze et al.,

2009) was launched in 2009, followed by the Orbiting Carbon Observatory 2 (OCO-2) (Crisp et al., 2004) in 2014. CO<sub>2</sub>-dedicated missions in development are: the Chinese Carbon Dioxide Observing Satellite, TanSat; the continuation of the GOSAT mission, GOSAT-2; the French Space agency's MicroCarb mission. The Active Sensing of CO<sub>2</sub> Emissions over Nights, Days, and Seasons (ASCENDS) mission will use lidar methods to perform night measurements as well (CEOS, 2014).

Inversion studies have exploited CO<sub>2</sub> retrievals to infer surface fluxes (Chevallier, 2015; Chevallier et al., 2005, 2009; Deng et al., 2014; Feng et al., 2017; Nassar et al., 2011; Reuter et al., 2014; Takagi et al., 2014) and the potential of these data to improve flux estimates through atmospheric inversion has been acknowledged. Nevertheless, current CO<sub>2</sub> retrievals are not accurate enough (Chevallier et al., 2014; Takagi et al., 2014) to obtain realistic, consistent flux estimates, so that inversion flux estimates still rely on in situ data.

### Transport model

Back to Figure 1.1, the function  $H$  that projects the fluxes into to observation space is represented by a numerical model of atmospheric transport—in this case, a set of equations that describe the actual atmospheric transport. The model solves for the large-scale processes of advection and horizontal diffusion that transport the tracers from zones where they are produced. Other transport processes, like turbulence and moist convection, take place at sub-grid scale and cannot be represented directly. Their mean effect on the tracer concentration at grid-scale is thus included through simplified representations called parameterizations. In principle, the transport of CO<sub>2</sub> is linear. But the discretization of the equations describing the transport introduces non-linearities in the function  $H$ . These non-linearities must be taken into account when solving for the inverse problem.

In some numerical models, the transport fields—meteorological fields of vertical and horizontal winds, moisture, temperature, pressure, etc—required to simulate tracer transport are calculated on-line, i.e. in step, with the transport calculations. On the other hand, in so-called off-line models the transport calculations are separated from the computationally-demanding dynamical calculations that generate the meteorology. Therefore off-line models run faster; they are fed with meteorological fields previously generated with an atmospheric meteorological model. Yet there are two limitations in the off-line approach. Meteorological data for transport processes at sub-grid scale that are generated in the on-line models are often unavailable for off-line models; instead the latter may use average values. Also off-line models use meteorological fields that are updated at intervals larger than in the atmospheric model, therefore at coarser temporal resolution, and also likely spatial, resolution.

There is a considerable spread in the concentrations estimated from the different transport models available (Geels et al., 2007; Law et al., 2008). The spread stems from models' biases in the representation of transport. Spatial resolution is also crucial to transport simulations. State-of-the-art global transport models achieve horizontal resolutions of up to 1 – 2° (Law et al., 2008) and up to 60 vertical levels (Law et al., 2008). Some global models have zooming capabilities that allow refining the resolution on a specific zone, from continental to regional scales. Regional transport models further refine the spatial resolution down to a few kilometers (Menut et al., 2013; Moreira et al., 2013), but in this case they need a field of CO<sub>2</sub> as lateral boundary conditions, which originates from a coarser-resolution model.

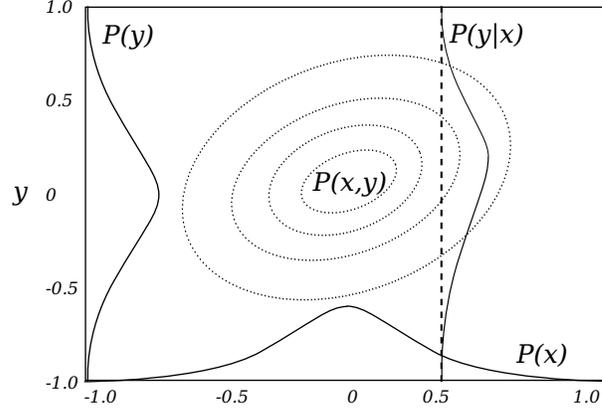


Figure 1.7: Illustration of the Bayes' theorem for the two-dimensional case. Adapted from [Rodgers \(2000\)](#).

### 1.1.2 Bayesian approach to the inverse problem

The Bayes' theorem provides the formalism usually adopted to solve for the inverse problem. Its application will be illustrated first in the case of scalars, and then generalized to the vector case. In Figure 1.1 a point in flux space  $x_b$  bearing an uncertainty  $\sigma_b$  is equivalent to formulate that there is a probability density function (PDF) with mean  $x_b$  and standard deviation  $\sigma_b$ . The same statement applies to an observation with errors. This way, the model projects the prior flux PDF into the observation space. After inversion, the optimized flux can be also described by a PDF with mean  $x_a$  and uncertainty  $\sigma_a$ . Its projection into the observation space would lie in the intersection of the PDFs of the prior and the observation. The interest lies in describing the intersection of those PDFs. Figure 1.7 illustrates a two-dimensional space defined by scalars  $x$  and  $y$ .  $P(x)$  represents the probability of  $x$  being in  $(x + dx)$ . Similarly,  $P(y)$  represents the probability that  $y$  lies in  $(y + dy)$ . Contours represent  $P(x,y)$ , the joint probability of  $x$  being in  $(x + dx)$  and  $y$  being in  $(y + dy)$ . If one had some prior estimate of  $x$  then one could define  $P(y|x)$  as the probability of  $y$  being in  $(y + dy)$  for a given value  $x$ .

From Figure 1.7 we can see that  $P(x)$  can be obtained by integrating  $P(x,y)$  along all the values of  $y$

$$P(x) = \int_{-\infty}^{\infty} P(x,y)dy \quad (1.2)$$

$P(y)$  can be calculated by integrating  $P(x,y)$  along all the values of  $x$ .  $P(y|x)$  is proportional to  $P(x,y)$  as a function of  $x$ . Since  $\int P(y|x) = 1$ , then it follows

$$P(x|y) = \frac{P(x,y)}{\int P(x,y)dx} \quad (1.3)$$

Substituting in the previous equation, we can formulate  $P(x)$  as

$$P(x) = \frac{P(x,y)}{P(y|x)} \quad (1.4)$$

Similarly, it can be shown that

$$P(y) = \frac{P(x,y)}{P(x|y)} \quad (1.5)$$

Combining both equations, we obtain

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \quad (1.6)$$

Equation (1.6) is the Bayes' theorem. The left-hand term is the posterior PDF of  $x$ , which denotes the updated prior estimate of  $x$  based on the information contained in the observation.  $P(y)$  is independent of  $x$  and is usually taken as a normalizing factor.

Assuming Gaussian, unbiased distributions of the errors in  $x$  and  $y$  (Tarantola, 2005), one can write:

$$P(x) = \frac{1}{\sigma_b \sqrt{2\pi}} \exp\left[-\frac{(x-x_b)^2}{2\sigma_b^2}\right] \quad (1.7)$$

where  $x_b$  is the prior estimate of  $x$  with an error  $\sigma_b$ . In the case of the observation, recalling the link between the fluxes and the observations, provided by the model  $H$ , we have

$$P(y|x) = \frac{1}{\sigma_o \sqrt{2\pi}} \exp\left[-\frac{(y-H(x))^2}{2\sigma_o^2}\right] \quad (1.8)$$

where  $\sigma_o$  is the observation error. Substituting these two equations in the Bayes' theorem and ignoring the normalizing factor one obtains:

$$P(x|y) \sim \exp\left[-\frac{(x-x_b)^2}{2\sigma_b^2} - \frac{(y-H(x))^2}{2\sigma_o^2}\right] \quad (1.9)$$

In the Bayesian framework, the optimal solution for  $x$  is the one that maximizes the posterior PDF. The optimal value  $x_a$  can be obtained by finding the minimum of the cost function

$$J(x) = \frac{(x-x_b)^2}{\sigma_b^2} + \frac{(y-H(x))^2}{\sigma_o^2} \quad (1.10)$$

Generalizing this to the vector case, we obtain

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y}_o - H(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \quad (1.11)$$

where  $\mathbf{x}_b$  is the control vector gathering the parameters (i.e. fluxes) controlled by the inversion and  $\mathbf{y}_o$  represents the set of observations, respectively. The errors in  $\mathbf{x}_b$  and  $\mathbf{y}_o$  are organized in matrices, and since they may be correlated,  $\mathbf{B}$  and  $\mathbf{R}$  are called the variance-covariance matrices of the errors in the prior and in the observations, respectively.

Assuming the transport model is linear (see next Section), i.e., it can be represented as a matrix, Equation 1.11 can be written as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(\mathbf{y}_o - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathbf{H}\mathbf{x}) \quad (1.12)$$

The set of optimal fluxes  $\mathbf{x}_a$  is the one for which the gradient of  $J(\mathbf{x})$  equals zero, and is given by

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y}_o - \mathbf{H}\mathbf{x}_b) \quad (1.13)$$

This equation can be written as

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y}_o - \mathbf{H}\mathbf{x}_b) \quad (1.14)$$

where  $\mathbf{K}$  is called the gain, or weight, matrix

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (1.15)$$

or alternative as (Bouttier and Courtier, 2002)

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{R}^{-1} \quad (1.16)$$

As explained in the next section, how the gain matrix is formulated has an important implication on the numerical solution of the inverse problem.

Note that (1.13) gives the statistically optimal fluxes we look for. But the complete solution to the inverse problem is a PDF, also Gaussian and unbiased, with expected value  $\mathbf{x}_a$  and covariance matrix  $\mathbf{A}$  given by

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} \quad (1.17)$$

Or alternatively,

$$\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1} \quad (1.18)$$

As with the gain matrix, the choice between (1.17) or (1.18) will depend on the dimensions of the inversion problem (see next section).

### 1.1.3 Solution of the inverse problem

#### Analytical method

In the analytical method, the control vector does not involve directly the fluxes. Instead the elements of  $\mathbf{x}$  are scaling factors that are applied to the fluxes. Thus each scaling factor corresponds to the flux budget for a given geographic area, period and flux type. The flux budgets are contained in the matrix  $\mathbf{H}$ .

$\mathbf{H}$  can be seen as the combination of three operators (Wu et al., 2016):

$$\mathbf{H} = \mathbf{H}_{\text{sample}}\mathbf{H}_{\text{tran}}\mathbf{H}_{\text{dist}} \quad (1.19)$$

The first operator,  $\mathbf{H}_{\text{dist}}$ , maps each scaling factor in  $\mathbf{x}$  to CO<sub>2</sub> fluxes on the grid of the transport model. These fluxes are known as a basis function. The second operator,  $\mathbf{H}_{\text{tran}}$ , is an atmospheric transport model that maps the base functions to concentrations of CO<sub>2</sub>. Finally,  $\mathbf{H}_{\text{sample}}$  generates the elements of  $\mathbf{H}$  by sampling the output the transport model at the location and time of the observations in  $\mathbf{y}_o$ . The elements of  $\mathbf{H}$  are referred to as response functions. Note that there might be parameters not controlled by the inversion. Their response function is included in a vector  $\mathbf{y}_{\text{fix}}$ , and (1.14) can be written as

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y}_o - \mathbf{H}\mathbf{x}_b - \mathbf{y}_{\text{fix}}) \quad (1.20)$$

The analytical method allows obtaining the explicit values of  $\mathbf{x}_a$  and its uncertainties. The feasibility of the method depends on the size of the problem. Calculating the matrix  $\mathbf{H}$  implies a transport simulation for each base function. In (1.15) and (1.16), calculating the gain matrix implies the multiplication of matrices with dimensions of the control and observation vectors. Therefore, the approach is impractical when both the number of control parameters and the number of observations is very large.

Solving for  $\mathbf{x}_a$  depends on the formulation of  $\mathbf{K}$  according to (1.15) or (1.16). The choice depends on the number  $n$  of parameters in the control vector and the number of observations  $m$ . If (1.15) is chosen, the operation implies calculating a matrix of  $m \times m$  elements. This is feasible if the number of observations is small. On the other hand, with (1.16), calculating  $\mathbf{K}$ , a matrix of size  $n \times n$ , is viable if the number of control parameters is small. For very large problems, the variational approach is more convenient.

### Variational method

Instead of calculating the gain matrix explicitly, the variational approach solves for the optimization problem by searching for an approximate optimal solution in an iterative manner (Bouttier and Courtier, 2002). The method approaches the minimum of the cost function  $J(\mathbf{x})$  using a descent algorithm, evaluating iteratively the cost function and its gradient

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y}_o - H(\mathbf{x})) \quad (1.21)$$

This assumes the forward model  $H$  is linear. However, the discretization of the transport equations introduces some non-linearities. The approach is to linearize  $H$  about a point  $\mathbf{x}_0$  so that  $H(\mathbf{x}) - H(\mathbf{x}_0) \approx \mathbf{H}(\mathbf{x} - \mathbf{x}_0)$ .  $\mathbf{H}$  is called the tangent linear of  $H$  about  $\mathbf{x}_0$ . The minimization process stops either when a limiting number of iterations has been achieved, or when the norm of the gradient reaches a predefined threshold.

## 1.2 Estimates of CO<sub>2</sub> fluxes in Amazonia from atmospheric inversion

Data paucity is limitation of global inversion systems to infer surface fluxes. Most sampling sites across the globe have been chosen to detect large-scale signals and avoid the influence of local fluxes; remote locations on land and ocean have been usually preferred for this purpose, leaving continental areas, like Amazonia, mostly undersampled. The most recent inversion intercomparison study from (Peylin et al., 2013) gathered results from 11 global inversion systems assessed in the frame of the RECCAP initiative (Regional Carbon Cycle Assessment and Process; Canadell et al., 2011). With different monitoring network configurations among them, none of the systems assimilated observations within tropical South America, much less in Amazonia, as shown in Figure 1.8.

Typically, global inversion results vary widely across Amazonia. Differences in transport models, prior fluxes, spatial and temporal discretization of the control parameters, observation network configuration and the inversion algorithm, to name a few key system attributes, contribute to the spread among their results, basically because there are no data within the Amazon basin.

Figure 1.9 shows the spatial distribution of mean annual land CO<sub>2</sub> fluxes (i.e., excluding fossil fuel emissions) for year 2000 from 8 global inversion models gathered by Peylin

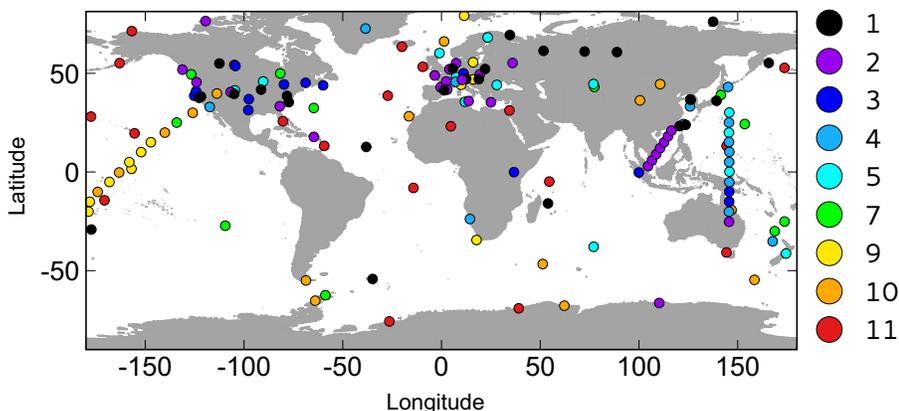


Figure 1.8: Map of site locations used by any of the inversions assessed by [Peylin et al. \(2013\)](#). Colors represent the number of inversion systems that use that site. Adapted from [Peylin et al. \(2013\)](#).

[et al. \(2013\)](#). The posterior flux patterns illustrate that in some inversions Amazon fluxes were optimized using a few scaling factors (e.g. Figure 1.9a), or even a single adjustment factor (e.g. Figure 1.9h). Other inversions used regular horizontal grid-cells (e.g. Figure 1.9e), with several cells within the Amazon basin, which provides more information about the potential heterogeneity of Amazon fluxes than inversions where Amazon fluxes were controlled using one single scaling factor.

Differences in inversion fluxes are reflected in the inter-annual variations as well. Figure 1.10 illustrates the time series of annual natural flux (excluding fire emissions) in tropical South America, a region containing mainly the Amazon basin, using Transcom’s region definition, from the 11 inversion compared in [Peylin et al. \(2013\)](#). In Figure 1.10 there exists a large spread among the inversion estimates, and inversions disagree on the magnitude of the flux anomalies during specific extreme climatic events that have affected Amazonia, such as the droughts of 2005 and 2010.

There exists a large spread in the mean seasonal cycle of NEE over the tropical South America region as well. Figure 1.11 shows the mean monthly NEE averaged over the period 2000 – 2008 for 7 inversions gathered by [Peylin et al. \(2013\)](#). In general, inversions indicate that vegetation have a peak of CO<sub>2</sub> uptake after the onset of the dry season in most of the Amazon basin, between April and May, associated with the end of the South American Monsoon System (SAMS). CO<sub>2</sub> uptake decreases during the transition between dry and wet seasons, between September and October, and the region turns into a source of CO<sub>2</sub> during the wet season during November – March.

The lack of observations in the tropical South America region is at the origin of the seasonal and inter-annual differences of the NEE in (mainly) the Amazon region, as given by global inversions. To illustrate this, it is interesting to look at the results from the study of [Stephens et al. \(2007\)](#). They evaluated the fit of 12 inversions systems—used in the Transcom 3 Level 2 inter-comparison experiment—to the mean annual daytime CO<sub>2</sub> vertical gradient between 1 and 4 km, averaged over 12 locations. Each inversion system is represented with a different letter in the Figure 1.12.

The models’ mean annual land carbon flux over the period 1992 – 1996 is plotted as a function of their predicted CO<sub>2</sub> vertical gradient. [Stephens et al. \(2007\)](#) found that

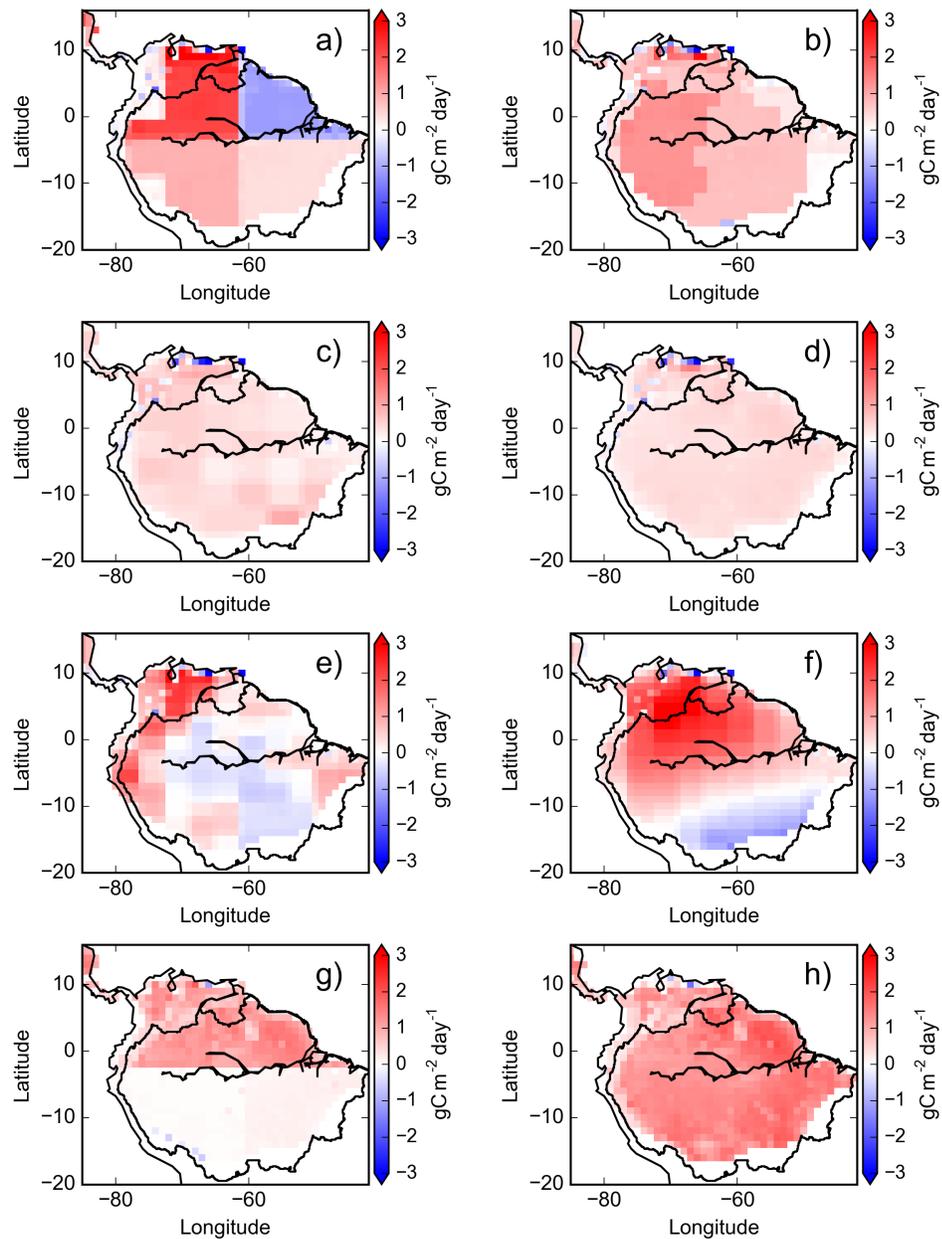


Figure 1.9: Annual mean natural CO<sub>2</sub> flux for tropical South America, as defined in the Transcom project, for 2000 for 8 inversion models submitted to RECCAP. Source: <http://transcom.lsce.ipsl.fr>.

the models that were consistent with the observed CO<sub>2</sub> vertical gradient also estimated a weaker Northern latitude sink, and a smaller tropical source. Since seasonal interactions between atmospheric mixing and NEE in the Northern Hemisphere create inter-hemispheric annual CO<sub>2</sub> gradients, and because the global mass balance of atmospheric carbon must be maintained, the inversion models introduce compensating variations in the

1. CO<sub>2</sub> EXCHANGES BY LAND ECOSYSTEM IN AMAZONIA INFERRED FROM ATMOSPHERIC INVERSION

---

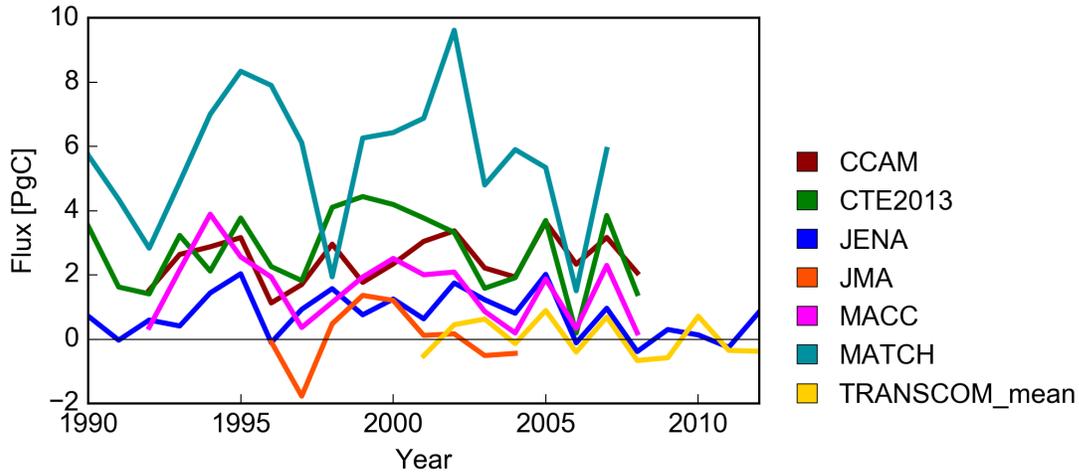


Figure 1.10: Annual natural CO<sub>2</sub> flux (excluding fire emissions) for tropical South America (Transcom definition) from 1990 to 2012 for the 7 inversion models submitted to RECCAP. Source: <http://www.globalcarbonatlas.org>.

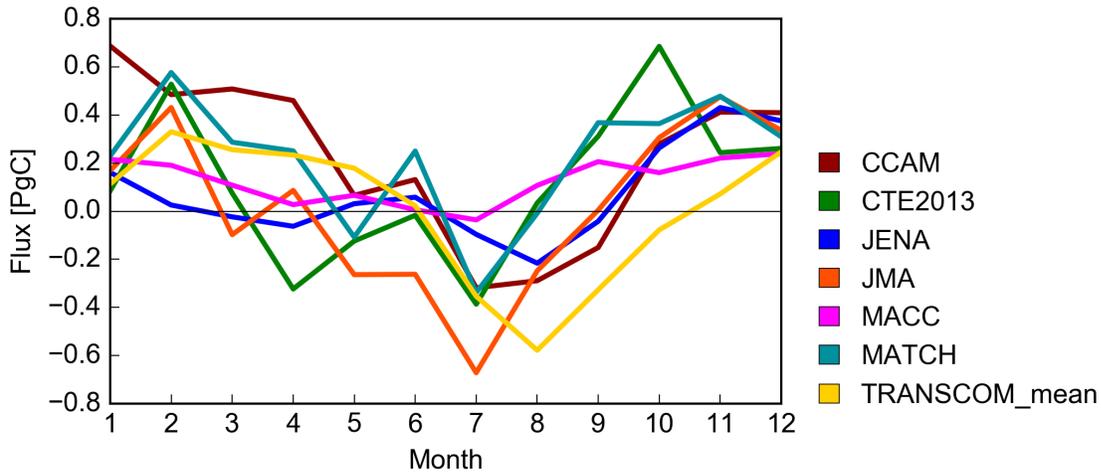


Figure 1.11: Mean monthly natural CO<sub>2</sub> flux for tropical South America (Transcom definition) from 2000 to 2008 for 7 inversion models submitted to RECCAP. Adapted from <http://www.globalcarbonatlas.org>.

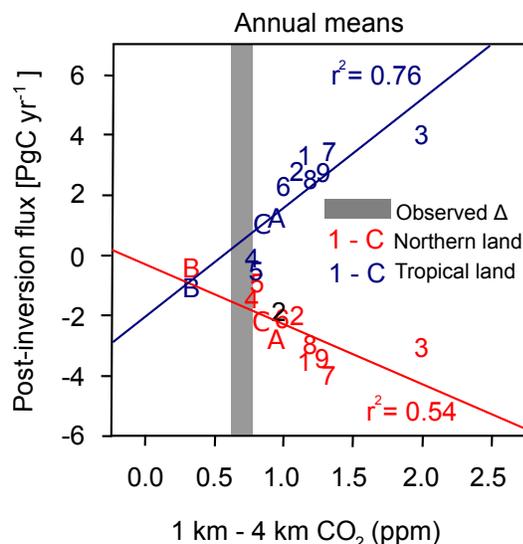


Figure 1.12: Mean annual Northern and tropical land carbon flux averaged over 1992 – 1996, estimated by the 12 inversion systems evaluated in 12 TRANSCOM3 Level 2 experiment, as a function of models' estimated mean vertical CO<sub>2</sub> gradient between 1 and 4 km. Vertical axis: estimated flux for northern (red) and tropical (blue) land regions. Horizontal axis: predicted 1 km - 4 km CO<sub>2</sub> gradient. Numbers (1 – 9) and letters (A – C) represent the models' estimates. Grey bar: represents the observed gradient (center) and its uncertainty (width). Adapted from [Stephens et al. \(2007\)](#).

tropical land fluxes. Thus, Figure 1.12 illustrates that systematic errors in the covariance between seasonal transport and NEE in the Northern Hemisphere induce an overestimation of the northern land sink, compensated by a stronger tropical land source. This means that in a region like Amazonia, with scarce observations of the atmosphere inside and around the region, local fluxes estimated through inversion are subject to misfits in remote sites distant from Amazonia.

In the mid-latitudes, the available potential energy associated with zonal temperature gradients is the main source of energy for synoptic variations. Latent heat and radiative heating play a secondary role in the energetics in these latitudes. In the tropics, however, latitudinal temperature gradients are small and latent heat release, mostly associated with convective cloud systems, appears to be the dominant energy source driving synoptic scale variations (p. 370 – 371; [Holton, 2004](#)). Thus, while synoptic variations outside the tropics are dominated by horizontal transport, vertical convective mixing prevails in the tropics. Accurate representation of these mechanisms in transport models is crucial in inverse modeling systems ([Gurney et al., 2003](#); [Parazoo et al., 2008](#)). Errors in the representation of vertical mixing induce biases in their flux estimates ([Stephens et al., 2007](#)).

### 1.3 Toward robust inverse modeling estimates of CO<sub>2</sub> fluxes in Amazonia

A few studies have used directly atmospheric CO<sub>2</sub> observations to estimate the Amazon CO<sub>2</sub> balance. Short-term aircraft campaigns with CO<sub>2</sub> measurements have used mass bal-

ance calculations of the atmospheric column to estimate regional CO<sub>2</sub> fluxes (Chou et al., 2002; Lloyd et al., 2007; Wofsy et al., 1988) that are representative of non-characterized area of central Amazonia. More regular aircraft profiles in the interior of the Amazon, completed by data from coastal sites, have allowed obtaining first estimates of seasonal CO<sub>2</sub> budgets for eastern Amazonia (Gatti et al., 2010). With a more complete data set of biweekly aircraft profiles, Gatti et al. (2014) (GA2014 henceforth) made an unprecedented effort to sample the atmosphere within the Amazon basin in such a way that their observations were sensitive to fluxes over most of the Amazon basin, covering not only a large fraction of Amazon forest areas but capturing the signal of savanna and agricultural lands.

GA2014 collected aircraft CO<sub>2</sub> data over two years with contrasting climatic conditions in Amazonia. In 2010 the region was anomalously dry for two reasons. During January – March, El Niño caused low precipitation in the northern and central Amazon basin. Then, in the second half of the year a positive anomaly in the North Atlantic sea surface temperature caused the inter-tropical convergence zone (ITCZ) to remain more northerly than usual. This prolonged and intensified the dry conditions in southern parts of the basin. 2011, on the other hand, was an anomalously wet year. With their observations and using a mass balance approach GA2014 calculated basin-wide fluxes over those two years. GA2014 found that during 2010 the Amazon emitted  $0.48 \pm 0.18$  PgC, but was carbon neutral during 2011 ( $0.06 \pm 0.1$  PgC yr<sup>-1</sup>). After removing CO<sub>2</sub> losses due to fire emissions, they found the Amazon basin was nearly neutral ( $-0.03 \pm 0.22$  PgC yr<sup>-1</sup>) during 2010, and that during the wet year the vegetation was a net sink of  $0.25 \pm 0.14$  PgC yr<sup>-1</sup>, consistent with a long-term estimate of  $0.39 \pm 0.10$  PgC yr<sup>-1</sup> from forest censuses. Thus, their results highlighted the importance of moisture conditions on the Amazon carbon balance. It was the first that such measurements were used to derive a basin-wide estimate of CO<sub>2</sub> fluxes in Amazonia.

van der Laan-Luijkx et al. (2015) were the first to assimilate the vertical profiles of GA2014 into a dedicated version of the CarbonTracker global inversion system (Peters et al., 2007). van der Laan-Luijkx et al. (2015) predicted that during 2010, year of the drought, the Amazon was net source of  $0.07 \pm 0.42$  to  $0.31 \pm 0.42$  PgC yr<sup>-1</sup>, the range being from a set of inversions using different biomass burning estimates. In contrast, during 2011 the ecosystem turned to net carbon sink of  $-0.15 \pm 0.42$  to  $-0.33 \pm 0.46$  PgC yr<sup>-1</sup>. These results were consistent with the mass balance approach of GA2014 but they gave a slightly smaller difference between 2010 and 2011 than GA2014. Alden et al. (2016) published results of the assimilation of airborne observations made over the longer period 2010 – 2012 into a regional inversion system, using two transport models. They predicted the Amazon basin to be a source of  $0.5 \pm 0.3$  PgC yr<sup>-1</sup> during 2010, in agreement with GA2014, but also estimated a net source of  $0.2 \pm 0.3$  PgC yr<sup>-1</sup> in 2011 (see Figure 3a in Alden et al. (2016)). With a longer observation record by one year, their results showed that during 2012 the vegetation continued to recover from the effects of drought and turned into a large net sink, with a net uptake larger by  $0.68 \pm 0.45$  PgC yr<sup>-1</sup> in 2012 with respect to 2010.

In line with the efforts of van der Laan-Luijkx et al. (2015) and Alden et al. (2016), the present work in this thesis aim at improving the knowledge of the carbon balance in Amazonia using an inverse modeling approach. By the time the present thesis works started in 2012, available data was much more limited with a few ground stations mainly near the coast. Some measurement sites had gaps of several years, and few sites were interrupted. With this limited data set, in the chapter *On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia* (Chapter 2), I analyzed the

seasonal and interannual variations of NEE in Amazonia with two global inversions, assimilating the data from 4 stations around the Amazon in one of them. The study looked at the global inversion results with more detail than previous intercomparison studies and showed the impact of these scarce data on the optimized fluxes. Though, the study remarked the need for more data, and suggested that assumptions on the prior errors and aspects of representation of the transport in Amazonia should be revisited. In the chapter *Regional atmospheric modeling of CO<sub>2</sub> transport in Amazonia* (Chapter 3), I evaluated the potential improvement brought by a regional atmospheric model, relative to a global one (which assimilates weather data from satellite and stations), when both models were used to force an off-line CTM to simulate CO<sub>2</sub> fields in Amazonia. I investigated the sensitivity of the transport of CO<sub>2</sub> to different transport components, as a function of four different observation vectors—based on the vertical profiles of GA2014—to bring insight on the best strategy to assimilate aircraft measurements. Finally, in the chapter *Regional atmospheric inversion of CO<sub>2</sub> in Amazonia* (Chapter 4), I present the regional inversion system I developed, which take advantage of both transport configurations, to optimize NEE using the four observation vectors separately. The chapter brings further information on the consistency and robustness of Amazon NEE from assimilation of different observation types.

## 1.4 Structure of this work

This dissertation starts with an analysis of the seasonal and interannual variations of the net ecosystem exchange (NEE) in Amazonia over the period 2002 – 2010 (Chapter 2). It describes the setup of two global inversion configurations to investigate CO<sub>2</sub> fluxes over the Amazon region, one of which included few ground-based station data that were available before the study of GA2014 and that had not been exploited in previous global inverse modeling studies. As a first indicator of the confidence on the impact of assimilating these new data, the impact on the posterior concentration time series at the new sites (Section 2.2.1) was characterized, followed by the analysis of the impact on the fluxes (Section 2.2.2). The spatial distribution of flux increments by both inversions was analyzed as a measure of the confidence on the improvement of the simulated seasonal cycle and interannual variability (Section 2.2.2) of NEE in Amazonia, both at basin and sub-regional scale.

The data collected by GA2014 have filled a major information gap in inverse modeling studies over Amazonia. The combination of new data and modeling tools, motivated the second part of my research (Chapter 3), which aimed at evaluating whether CO<sub>2</sub> transport in Amazonia could be improved with the meteorology generated with the regional atmospheric model BRAMS, developed for simulating meteorology and chemistry over tropical South America, with respect the global forecast system ECMWF. Thus, two transport configurations were setup (Section 3.1.1). The meteorology of both models was evaluated against ground-based, remote-sensing and satellite observations (Section 3.2.1). The quality of the simulated CO<sub>2</sub> fields by both models is assessed (Section 3.2.2). With the data of GA2014, not only vertical profiles were evaluated. These data allowed evaluated the quality of the horizontal transport. This chapter gives an estimate of the uncertainty of CO<sub>2</sub> related to differences between the two transport models, which are prescribed with the same surface fluxes.

Chapter 4 describes the third part of this research. It describes the setup of two in-

version systems (Section 4.1) based on the forward transport configurations evaluated in Chapter 3. Individual inversions were performed with different observation vectors, including vertical profiles and different types of horizontal gradients, with both inversion systems. The study evaluated the fit to the observations (Section 4.2), and the impact on the optimized fluxes (Section 4.3). Finally, Chapter 5 presents a summary of the main results this thesis, as well as suggestions on the orientation of future research.

---

## On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia

When starting the thesis in 2012, there existed few sampling sites in South America with CO<sub>2</sub> measurements with sufficient precision to be used in atmospheric inversion studies. However, these data had not been assimilated in most global inversion systems (Peylin *et al.*, 2013). Motivated by this, I assimilated CO<sub>2</sub> data from some of the few ground-based stations were available at the time into a global inversion system, aiming at improving the knowledge on the seasonal and interannual variations of the net ecosystem exchange in Amazonia, and assess the impact of local observation sites in the global system. The results of this chapter lead to the publication "On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia" in the journal *Atmospheric Chemistry and Physics* (Molina *et al.*, 2015). The study is presented in Section 2.1, the main results and their implications are detailed in Section 2.2, and the concluding remarks in Section 2.3. The whole publication is given in the Appendix A.

### 2.1 Objective of the study

The study aimed at improving our knowledge of the seasonality and interannual variations of NEE in Amazonia over the period 2002 – 2010. This period is long enough to assess the impact of extreme climatic events such as the severe droughts of 2005 and 2010, and the extreme humid conditions registered in 2009. For this purpose, I designed an inversion, referred to as INVSAM henceforth, by building on the Monitoring of Atmospheric Climate and Composition (MACC) global inversion system version 10.1 (MACCv10.1 hereafter), described in detail by (Chevallier *et al.*, 2010). Both INVSAM and MACCv10.1 optimized a prior NEE estimate simulated with the ecosystem model ORCHIDEE by Maignan *et al.* (2011). Ocean fluxes were also optimized, but the study focused on the corrections to the NEE.

MACCv10.1 assimilated data from about 100 sampling sites around the globe, but none were located in Amazonia. To improve the NEE estimate, INVSAM added four local sites. Two sampling sites were located at Arembepe and Maxaranguape, on the eastern coast of Brazil, basically under marine influence, measured the background CO<sub>2</sub>. The other two sites at Santarém in northeastern Brazil, and in French Guiana, were surrounded by

tropical forests and captured the signal of terrestrial ecosystems. This way, the impact of local data on the optimized fluxes was assessed against MACCv10.1. In addition, the study went deeper than previous inversion intercomparison studies such as TRANSCOM and RECCAP (Canadell et al., 2011), since it evaluated the inversion results over and within Amazonia in an attempt to identify robust patterns.

## 2.2 Main results and implications

### 2.2.1 Simulated vs. observed concentrations

Overall, both INVSAm and MACCv10.1 significantly reduced the misfit to the observations with respect to the prior. However, misfits were further decreased with the assimilation of data in South America. Examination of simulated CO<sub>2</sub> time series at the local sites, showed that assimilating local data significantly rescaled the seasonal variations (e.g. at the site in French Guiana), and even changed the seasonality (e.g. at Santarém and Maxaranguape) in agreement with the observations, with respect to the reference inversion. Correlation between time series of optimized and observed daily concentrations were rather low in both inversions at the four local sites. However, correlations at monthly scale were higher with both inversion estimates, and again, INVSAm showed better correlation than MACCv10.1. This suggested a potential improvement of seasonal variations of NEE in Amazonia.

### 2.2.2 Impact on surface biogenic CO<sub>2</sub> fluxes

#### Spatial distribution of flux corrections

In terms of the spatial distribution of flux corrections, both INVSAm and MACCv10.1 predicted large continuous, nearly zonal patterns that extended over both land and ocean, likely due to the dominant atmospheric circulation. On the continent, both inversions predicted flux corrections of opposite sign between north and south of tropical South America, between austral summer and winter and at the annual scale. This result is typical of inversion systems over undersampled regions. The assimilation of the local stations, however, did increase the amplitude of the dipole. Flux corrections by INVSAm often exceeded 150% of the prior fluxes (in absolute value), indicative of significant flux corrections. In addition, with local data the inversion changed the latitudinal position of the dipole over the continent. This provides evidence of the large-extent impact of the information from local stations, despite the relatively small correlation scale length in prior uncertainties, and the limited area of the station footprints with high sensitivity to land fluxes.

#### Seasonality

At the scale of tropical South America, the seasonal cycle of optimized NEE in both INVSAm and MACCv10.1 seemed strongly driven by the prior NEE. Nevertheless, while flux corrections introduced by local data likely came from the signal of tropical broad-leaved (TBE) forests, the seasonal cycle could have remained driven by the NEE from other plant functional types (PFTs) represented in ORCHIDEE. Over TBE forests, however, prior and inversion estimates did not indicate clear correlation of NEE with rainfall or solar radiation. The results disagreed with the estimate of Jung et al. (2011). Prior and inverted NEE indicated local extremes, possible due to the overlapping of significantly different seasonal cycles from other subregions within Amazonia.

To examine if the inversions captured the potential spatial heterogeneity of NEE seasonality, two sub-regions were analyzed. Nevertheless, neither the prior nor the inversions indicated a clear seasonal cycle. Jung et al. (2011), on the other hand, continued to exhibit a smooth seasonal cycle, of nearly the same amplitude in both sub-regions and across all TBE forests. This reduced the confidence in Jung et al. (2011) to capture the spatial heterogeneity seasonal NEE variations. Within the two sub-regions, results suggested significant corrections to the seasonal cycle of NEE, especially over regions well constrained, in space and time, by the observation sites. However, large corrections were also applied over areas over which the local stations exhibited low sensitivity, particularly in central western Amazonia. This can be partly explained by the control exerted on the optimized fluxes by the local stations, as well as stations distant from Amazonia, suggested by the analysis of the spatial distribution of flux corrections. This is a direct consequence of the limited overlapping between measurement records at the local sites.

### Inter-annual variations

The analysis of the inversions' ability to capture the interannual variability of NEE within Amazonia focused on the drought periods recorded in 2005 and 2010, and the extreme humid conditions of 2009. The analysis was based on NEE anomalies with respect the mean NEE during 2002 – 2010 predicted by the different estimates. Thus, from the atmospheric perspective, negative anomalies were interpreted as an enhancement of the CO<sub>2</sub> sink, and positive anomalies as a reduction. I made an additional inversion using a "flat prior" (called FLAT), i.e., prior NEE without interannual variability with respect to the mean NEE 2002 – 2010 across tropical South America (see the publication in Appendix A for details). The goal was to assess the inversion system's ability to reproduce interannual variations from the observations alone, a potential indicator of the robustness of the inversion estimates.

Literature reports on the intensity and extension of the drought events of 2005 and 2010, and their negative effects on Amazonian forests. Nevertheless, across tropical South America prior NEE and both INVSAm and MACCv10.1 estimates indicated only small positive anomalies in those years, while FLAT predicted positive and negative anomalies in 2005 and 2010, respectively. In 2009, on the other hand, prior and inversion estimates all agreed on a large negative anomaly, which provided confidence on this pattern. Over TBE forests, prior and both INVSAm and MACCv10.1 predicted a positive anomaly in 2005 and negative in 2010. The latter contrasts not only with the anomalies inferred for the whole tropical South America, but also with the findings of Gatti et al. (2014). Their results, obtained using a mass balance approach and airborne CO<sub>2</sub> and CO observations, indicated CO<sub>2</sub> uptake reduction due to water stress during 2010 across Amazon forests. My results, therefore, are to be interpreted cautiously, as the small anomalies, compared to those considering all PFTs, suggest low NEE sensitivity in TBE forests to interannual climatic variations.

Interannual variations within the subregions examined in the analysis of seasonal variability revealed contrasting responses to drought events within and between the regions. While in eastern Amazonia the prior and both INVSAm and MACCv10.1 predicted positive and negative anomalies in 2005 and 2010, respectively, in central western Amazonia anomalies were negative in both years. Anomalies in J2011 were at least one order of magnitude smaller than the prior and both inversion estimates. Consequently, J2011 did not provide enough information on the heterogeneity of NEE interannual variations and help interpret my results.

## 2.3 Conclusions

In an effort to improve the knowledge on the seasonal and interannual variations of NEE in Amazonia I analyzed the results two global inversions, and assimilated data from four ground-level sites in tropical South America in one of the inversions. The results showed that information from both local and distant sampling sites controlled NEE in tropical South America. This control manifested as alternate zonal areas of positive and negative flux corrections over land and over large portion of the ocean in the Southern Hemisphere in a "dipole" pattern. The main effect of the local ground-based station data was to increase the amplitude of the dipole and change its position, and create some local patterns around the local sites.

In spite of the overall improvement of seasonal variations of the simulated CO<sub>2</sub> concentrations at the local sites, the seasonal cycle of NEE over tropical South America remained mostly unchanged in the inversion estimates. Particularly over rain forests, prior and inversion estimates disagreed with the assumption of higher CO<sub>2</sub> uptake during periods of higher solar radiation. The smooth patterns of NEE in [Jung et al. \(2011\)](#) disagreed with my results. The low reliability on the seasonal patterns of NEE from the inversions seems to confirm that the spatial distribution of flux corrections over the entire Amazon basin is an artifact from the inversion, reflecting the low density of the local monitoring network, and the limited measurement records.

Such considerations also reduce the confidence on the patterns of NEE interannual variability from the inversion INVSAm. Yet, some patterns were present across prior and inversion estimates, suggesting they are robust, such as the vegetation being a net source of CO<sub>2</sub> ( $\sim 0.21$  PgC) during 2005—year of the drought—and a strong sink ( $\sim -1.1$  PgC) during 2009, likely due to extreme rainfall conditions in 2009.

---

## Regional atmospheric modeling of CO<sub>2</sub> transport in Amazonia

The analysis in Chapter 2 reveals the limitations of the global-scale atmospheric inversion approach to estimate the CO<sub>2</sub> net ecosystem exchange (NEE) over the Amazon basin, even when assimilating regional CO<sub>2</sub> observations. In particular, it highlights the difficulties to simulate the concentrations at the measurement sites, especially near the coast, with a coarse resolution global-scale transport model like LMDZ. The representativeness and transport modeling errors of this global model are assumed to be two of the main explanations for the lack of ability to control the regional NEE in the inversion experiments presented in Chapter 2. In order to overcome such issues, I have developed regional and mesoscale atmospheric transport model configurations for the Amazon basin. This chapter presents these configurations and the evaluation of their skill for modeling the regional CO<sub>2</sub> transport while Chapter 4 will analyze the regional atmospheric inversions that I have conducted with these transport configurations.

When using regional transport model instead of global models, the simulations of the CO<sub>2</sub> transport can be improved both by increasing the spatial resolution and by using the meteorological forcing from regional meteorological models that are adapted to the specific conditions in the area of interest. The Brazilian developments on the Regional Atmospheric Modeling System (BRAMS; Freitas et al., 2009; Moreira et al., 2013) has been specifically developed to better represent meteorology and atmospheric chemistry-transport over South America. In Amazonia, vertical transport due to moist convection is a dominant process over advection, and largely explains synoptic variations of atmospheric CO<sub>2</sub> (Parazoo et al., 2008). BRAMS includes new physical schemes to better represent deep and shallow convection in the Amazon (Freitas et al., 2009). Another important source of uncertainty in model simulations is the strong dependency on the representation of surface processes. In this regard, BRAMS has been coupled "on-line" with the JULES land surface scheme, which has improved the representation of several meteorological variables (Moreira et al., 2013).

I prepared two regional transport configurations to simulate CO<sub>2</sub> transport over tropical South America at a high resolution of  $\sim 35$  km for the year 2010. I used the mesoscale and regional tracer-transport model (CTM) CHIMERE (Menut et al., 2013), driven with different meteorological fields provided by two atmospheric models: the integrated forecast system of the European Centre for Medium-Range Weather Forecasts (ECMWF), and

simulations with a configuration of the regional atmospheric model BRAMS that I developed for this study. I prepared the regional configuration and started the development of the interface between BRAMS and CHIMERE during a 4-month visit to the Brazilian Center for Weather Forecasts and Climate Studies (CPTEC) in March – June 2014, with the support of the Grupo de Modelagem da Atmosfera e Interfaces (GMAI). During that period, I launched the first BRAMS simulations and tests for output validation.

The evaluation of these two transport configurations focuses first on the meteorological forcing and on the potential improvement obtained by using the regional meteorological model BRAMS for the simulation of meteorological variables. I compared BRAMS and ECMWF to data from an extensive meteorological observation network in South America. Next, in order to check whether improvements of the regional meteorological forcing led to an improvement of the simulations of CO<sub>2</sub> transport in the region, I compared the simulated CO<sub>2</sub> fields using CHIMERE driven with both meteorological forcings, but prescribing the same set of surface CO<sub>2</sub> fluxes to CO<sub>2</sub> measurements in the Amazon basin.

As to transport, airborne CO<sub>2</sub> measurements provide essential information about the vertical and horizontal CO<sub>2</sub> distribution to validate and improve transport models (Gloor et al., 2000). Several aircraft campaigns collected CO<sub>2</sub> measurements across Amazonia: the Amazon Boundary Layer Experiment ABLE-A (Wofsy et al., 1988), ABLE-2B (Chou et al., 2002), the LBA Airborne Regional Experiment (LBA-CLAIRE) (Lloyd et al., 2007), the Balanço Atmosférico Regional de Carbono na Amazônia (BARCA) project (Andreae et al., 2012), and regular biweekly observations established by Gatti et al. (2010) and Gatti et al. (2014) (GA2014 henceforth). Some of these studies have served as data sources for regional transport model validation for CO<sub>2</sub> (Moreira et al., 2013), aerosols (Rosário et al., 2013), and other chemical species (Andreae et al., 2012; Bela et al., 2015). The airborne dataset of GA2014 are thus used to validate the regional CO<sub>2</sub> transport simulation in this study. Furthermore, the airborne dataset of GA2014 can also be considered as the most extensive regional CO<sub>2</sub> dataset for the atmospheric inversions in the Amazon basin. The NEE inversions of studies of van der Laan-Luijkx et al. (2015) and Alden et al. (2016) (Chapter 1) used these airborne measurements. Therefore, the regional atmospheric inversion framework I developed in this PhD thesis aimed at assimilating these data (see Chapter 4) and the main objective of the evaluation of the transport model in this chapter was to provide a basis for the set-up of the assimilation of these airborne data.

In particular, both van der Laan-Luijkx et al. (2015) and Alden et al. (2016) assimilated individual measurements of each vertical profile from GA2014. But this is not necessarily the best strategy to exploit aircraft data to constrain NEE. In particular, the assimilation of CO<sub>2</sub> spatial gradients from such data (Bréon et al., 2015) may increase the robustness of the constraint on the net ecosystem exchange (NEE) while reducing the sensitivity of the inversion to uncertainties in the signal of other regional and remote flux components. Therefore, this chapter investigates the ability to filter information on the NEE, and to separate it from that from other types of fluxes (from biomass burning, fossil fuel combustion, ocean exchanges and remote natural and anthropogenic activities) from the aircraft measurements or from spatial gradients between these measurements. This investigation is based on the analysis of the individual impacts from the different types of fluxes on the aircraft profiles and on sensitivity experiments with both transport configurations, where the estimates for different type of fluxes are compared to reference experiments. I investigated the wet-to-dry (April – May) and the dry-to-wet (October – November) transitions seasons in 2010—two periods with different meteorological regimes and highly uncertain in terms of NEE (Chapter 2).

In the following, Section 3.1.1 presents the transport configuration, the methods for validating the simulated meteorological fields and for assessing the CO<sub>2</sub> distribution. Section 3.2 details the comparison of model simulations and meteorological observations, as well as the comparison of modeled CO<sub>2</sub> and aircraft observations and spatial gradients. The results are discussed in Section 3.3 and the conclusions are presented in Section 3.4.

## 3.1 Materials and methods

### 3.1.1 Transport configuration

In this chapter, tropical South America was defined as the geographic area between 85.1° W – 19.7° W and 16.3° N – 23.1° S (Figure 3.1). Transport was simulated with the tracer transport model CHIMERE (Menut et al., 2013). CHIMERE also includes a package for chemistry, which was de-activated for the transport of CO<sub>2</sub>, an inert gas. Using meteorological fields, estimates of emissions, and initial and boundary conditions of CO<sub>2</sub> at the lateral and top boundaries of the simulation domain, CHIMERE predicts the concentration fields of CO<sub>2</sub> at mesoscale. In this study, CHIMERE was configured with a horizontal resolution of ~ 35 km, and 29 vertical levels from the surface up to an altitude of approximately 10 km (~ 300 hPa). Table 3.1 summarizes the components of the transport simulations. The table gives the reference of different estimates of NEE fluxes, atmospheric CO<sub>2</sub> boundary conditions and meteorology that were used in this chapter. Among the multiple model options, transport by deep convection was activated or switched off for additional sensitivity tests of the CO<sub>2</sub> concentrations to convective transport.

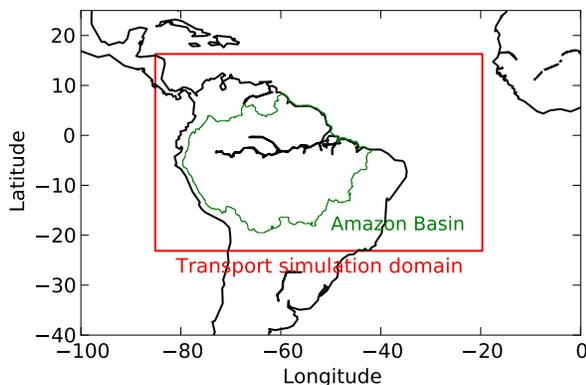


Figure 3.1: Geographic domain for CO<sub>2</sub> transport simulations with the offline transport model CHIMERE.

### Meteorological forcing fields from ECMWF

A first transport configuration, CH-ECMWF, was set up by forcing CHIMERE with the 3D meteorological fields of the ECMWF operational forecast. The relevant meteorological fields to simulate CO<sub>2</sub> transport in CHIMERE, listed in Table 3.2, were obtained at ~ 15 km horizontal resolution, with a 3-hour time step. CHIMERE's interface to ECMWF was used as is to interpolate the input meteorology to CHIMERE's horizontal resolution with hourly

frequency. Vertical interpolation to the transport model's vertical grid was performed by CHIMERE's meteorological preprocessor.

Table 3.1: Components of the transport simulations.

Component	Description
Atmospheric forcing	
BRAMS	Atmospheric fields at $\sim 35$ km
ECMWF	Atmospheric fields at $\sim 15$ km
Biosphere fluxes	
ORCHIDEE	Net ecosystem exchange (NEE) from ORCHIDEE simulation for 2010 at $2^\circ$ resolution and 3-hour time step, forced with CRUNCEP reanalysis data.
MACC	Optimized NEE from the Monitoring Atmospheric Climate and Composition v12.3 product (assimilation of global network), covering 35 years (1979–2013), at 3-hour time step and $3.75^\circ \times 1.88^\circ$ resolution.
MACC+	Optimized NEE from an inversion using the setup of MACC, assimilating the global network plus the observations of Gatti et al. (2014). Only observations within the first 6 levels of the transport model are assimilated (approximately 1500 m from the surface. The optimization is only for the period 2010–2011. Horizontal resolution $3.75^\circ \times 1.88^\circ$ and 3-hour time step.
Biomass burning CO <sub>2</sub> emissions	
GFED3.1	Monthly mean emissions from GFED version 3.1 at $0.5^\circ$ resolution.
Fossil fuel CO <sub>2</sub> emissions	
EDGAR4.2	Annual emissions from EDGAR 4.2 FT database at $0.1^\circ$ resolution.
Air-sea CO <sub>2</sub> exchange	
TAKA09	Annual mean flux from <a href="#">Takahashi et al. (2009)</a> at $3^\circ \times 2^\circ$ resolution.
CO <sub>2</sub> boundary conditions	
BC_MACC+	Optimized concentration fields from MACC+.
BC_CTracker	Optimized concentration fields from CarbonTracker-TM5 ( <a href="#">Krol et al., 2005</a> ; <a href="#">Peters et al., 2007</a> ).
Land use CO <sub>2</sub> emissions	Not included.

Table 3.2: List of meteorological fields to force the chemistry-transport model CHIMERE for simulation of CO<sub>2</sub> transport.

Variable
Zonal wind
Meridional wind
Air temperature
Specific humidity
Atmospheric pressure
Air temperature at 2 m
Planet boundary layer height
Surface sensible heat flux
Surface latent heat flux
Wind speed at 10 m
Topography
Friction velocity
Convective fluxes (entrainment and detrainment in updraft and downdraft)

### Meteorological forcing fields from BRAMS

In the second transport configuration, CH-BRAMS, CHIMERE was driven with meteorological fields (see Table 3.2) generated by a specific simulation of BRAMS with a frequency of 3 hours at the same spatial resolution than CHIMERE (i.e.  $\sim 35$  km); thus, no spatial interpolation was required even though the development of specific pre-processing tools was needed to convert BRAMS output into the relevant input variables for CHIMERE. BRAMS is described here in more detail since I constructed a dedicated configuration of BRAMS for tropical South America to calculate the meteorological fields for setting my transport simulations.

BRAMS is the atmospheric physical component of a coupled system which includes a chemistry-transport component and a land-surface scheme. Chemistry and transport routines are grouped in the Chemistry Coupled Aerosol and Tracer Transport model (CCATT) (Freitas et al., 2009; Longo et al., 2013). The exchanges of heat, moisture, surface momentum, radiation and trace gases between the surface and the atmosphere are calculated with the JULES land surface scheme (Clark et al., 2011; Best et al., 2011). The three models have been coupled on-line in the JULES-CCATT-BRAMS system (Moreira et al., 2013). BRAMS requires topography and meteorological initial and boundary conditions. Following (Moreira et al., 2013), the topography was prescribed with the 1-km resolution data set described by (Gesch et al., 1999). Atmospheric initial and boundary conditions were obtained from ECMWF reanalysis ERA-Interim (Berrisford et al., 2009) with a horizontal resolution of  $0.25^\circ$  and 26 vertical pressure levels at 6-hour intervals, interpolated on the transport model grid.

JULES requires grid-based data of vegetation cover and soil type, the normalized difference vegetation index (NDVI), sea surface temperature, soil carbon, air moisture and temperature. Following (Moreira et al., 2013), vegetation cover was prescribed using data from the PROVEG project (Sestini et al., 2003) within the Amazon Basin, and the vegetation map of Olson (1999). Soil texture was taken from the RADAMBRASIL project (Rossato

et al., 1998) for Brazil and from the FAO outside the country (Zobler, 1999). JULES uses NDVI to initialize the values of leaf area index and surface albedo. NDVI is obtained from 16-day composites derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) data from the period 2001-2002. Weekly mean sea surface temperature was obtained from the National Centers for Environmental Prediction (NCEP) (Reynolds et al., 2002). Soil carbon and moisture reach equilibrium over long timescales. A technique called spin-up (Yang et al., 1995) is normally used to bring soil carbon and moisture to equilibrium, but it is a computationally expensive process because it requires long simulations with the model. Therefore, initial soil carbon content was provided through values observed during the Large Scale Biosphere-Atmosphere Experiment in Amazonia (LBA) project (Batjes, 1996). The operational product developed by (Gevaerd and Freitas, 2006) was used to initialize the soil moisture content. Initial soil temperature was provided by the first level of BRAMS. CO<sub>2</sub> does not intervene in atmospheric dynamics. However, in the coupled system, JULES requires mean CO<sub>2</sub> concentration to simulate photosynthesis coupled with heat, water and momentum fluxes that influence atmospheric dynamics. Therefore, CO<sub>2</sub> initial and boundary conditions for photosynthesis simulation in JULES were obtained from 3-hour mean mole fractions from the CarbonTracker-TM5 (CarbonTracker, henceforth) system (Peters et al., 2007; Krol et al., 2005) with  $3^\circ \times 2^\circ$  horizontal resolution and 32 vertical levels, interpolated to my regional model grid. Note that although CCATT transports CO<sub>2</sub> based on the input from JULES and BRAMS, I used only the output of the atmospheric component of this coupled system to force CHIMERE. Thus, I controlled the input fluxes to my regional simulation for the analyses and for the inversion in Chapter 4.

#### CO<sub>2</sub> initial and lateral boundary conditions

In this study, for both transport configurations, CO<sub>2</sub> initial and boundary conditions were provided through post-inversion CO<sub>2</sub> fields from a global atmospheric inversion, based on the version 12.3 of the Monitoring Atmospheric Climate and Composition system (MACC) initially described by Chevallier et al. (2010). This inversion, designated MACC+, covered the period 2010 – 2011 and assimilated the observations from aircraft data of GA2014 (see Section 3.1.3 for details about these data) in addition to the global observation network data assimilated in MACCv12.3. In MACC+, only aircraft observations within the first 6 vertical levels of LMDZ, the transport model used in MACC, were assimilated. The optimized CO<sub>2</sub> fields, initially provided every 3 hours at  $3.75^\circ \times 1.88^\circ$  horizontal resolution and 39 vertical levels, were interpolated to CHIMERE's horizontal resolution to define the initial and boundary conditions called BC-MACC+.

In the sensitivity tests described in Section 3.2.2, CO<sub>2</sub> transport was simulated also using another CO<sub>2</sub> field for CO<sub>2</sub> initial and boundary conditions, namely from CarbonTracker, for CO<sub>2</sub> initial and boundary conditions defined BC-CTracker.

#### Surface CO<sub>2</sub> fluxes

Estimates of NEE, CO<sub>2</sub> release from biomass and fossil fuel burning, and air-sea CO<sub>2</sub> exchanges provided to CHIMERE are given in Table 3.1. In the reference experiments, NEE, was obtained from a simulation of the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) model (Krinner et al., 2005) at 3-hour and  $2^\circ$  resolution, forced with meteorological fields from CRUNCEP reanalysis (Sitch et al., 2015). These fluxes are referred to as ORCHIDEE henceforth. CO<sub>2</sub> fluxes from this simulation were

annually balanced and did not account for land-use change emissions or fires, but they resolved the diurnal cycle of NEE. In addition, 3-hourly NEE, optimized every 8 days from the MACCv12.3 and MACC+ global inversions (Table 3.1) were also used in sensitivity experiments (Section 3.2.2).

Biomass burning CO<sub>2</sub> emissions were taken from the Global Fire Emissions Database, version 3.1 (Randerson et al., 2013), which provided monthly emissions at 0.50° resolution. Fossil fuel CO<sub>2</sub> emissions were provided at 1-year temporal resolution by EDGAR4.2 FastTrack 2010 database (EU-JRC/PBL, 2013), scaled with annual global totals from the Global Carbon Project 2013 (Boden et al., 2013). Ocean fluxes were taken from the monthly climatology of CO<sub>2</sub> air-sea exchange from Takahashi et al. (2009). Land-use change (LUC) CO<sub>2</sub> emissions are prominent in the Amazon basin (~0.2 PgC yr<sup>-1</sup> between 2000 and 2010; Song et al., 2015), but LUC was not simulated by ORCHIDEE. Nevertheless, LUC CO<sub>2</sub> emissions were implicit in the optimized NEE from MACCv12.3 and MACC+ (ORCHIDEE-MACC and ORCHIDEE-MACC+, henceforth). Note that GFED3.1 included emissions associated to deforestation fires, which are a fraction of the total LUC CO<sub>2</sub> emissions.

Figure 3.2 shows the mean seasonal distribution of total CO<sub>2</sub> fluxes, combining NEE, fire and fossil fuel CO<sub>2</sub> emissions. Fluxes were averaged during the dry (July – October) and wet (November – June) periods of 2010. During the dry period, the three NEE estimates predict CO<sub>2</sub> emissions in areas dominated by savanna vegetation and in forest-savanna transition areas (Figure 3.5), along the southeastern and southern borders of the Amazon basin. Over these areas, ORCHIDEE predicted the highest emissions, while ORCHIDEE-MACC showed that assimilation of CO<sub>2</sub> data from the global network reduced the emissions. Assimilating local aircraft profiles in ORCHIDEE-MACC+ reduced emission further (Figure 3.14c). In Amazon forests ORCHIDEE simulated low CO<sub>2</sub> uptake over northern Amazon, and low CO<sub>2</sub> release in the area in the middle of the profile sites. ORCHIDEE-MACC and ORCHIDEE-MACC+, however, reverted that pattern and forests across northern Amazon were a source of CO<sub>2</sub>—similar to MACCv10.1 and INVSAm in Chapter 2, while the areas in middle and around the profile sites turned into a CO<sub>2</sub> sink. The assimilation of the local sites produced a slightly stronger uptake around the local sites—unlike INVSAm in Chapter 2, which produced stronger release—and shifted the uptake area to the east.

During the wet period, NEE from ORCHIDEE in the savanna areas was a strong CO<sub>2</sub> sink, more intense than in forest areas within the basin (Figure 3.2d). In ORCHIDEE-MACC and ORCHIDEE-MACC+ the uptake was weaker than ORCHIDEE over savanna areas. Note also that ORCHIDEE predicted CO<sub>2</sub> uptake across most of tropical South America. But ORCHIDEE-MACC predicted the land north of 10° S was a source of CO<sub>2</sub>, producing a dipole pattern (Figure 3.2e; see also Chapter 2). With the assimilation of local sites, ORCHIDEE-MACC+ predicted CO<sub>2</sub> uptake across northern and northeastern Amazon forests, and an area of CO<sub>2</sub> release east of ALF and SAN.

### 3.1.2 Evaluation of simulated meteorological fields

The quality of the simulated meteorology should be a first indicator of a model’s ability to simulate CO<sub>2</sub> transport. The simulated meteorological fields were confronted to vertical profiles from radiosonde observations at 11 sites distributed across the Amazon Basin (Figure 3.3), obtained from the National Oceanic and Atmospheric Administration – Earth System Research Laboratory (NOAA/ESRL) Radiosonde Database. These sites are usu-

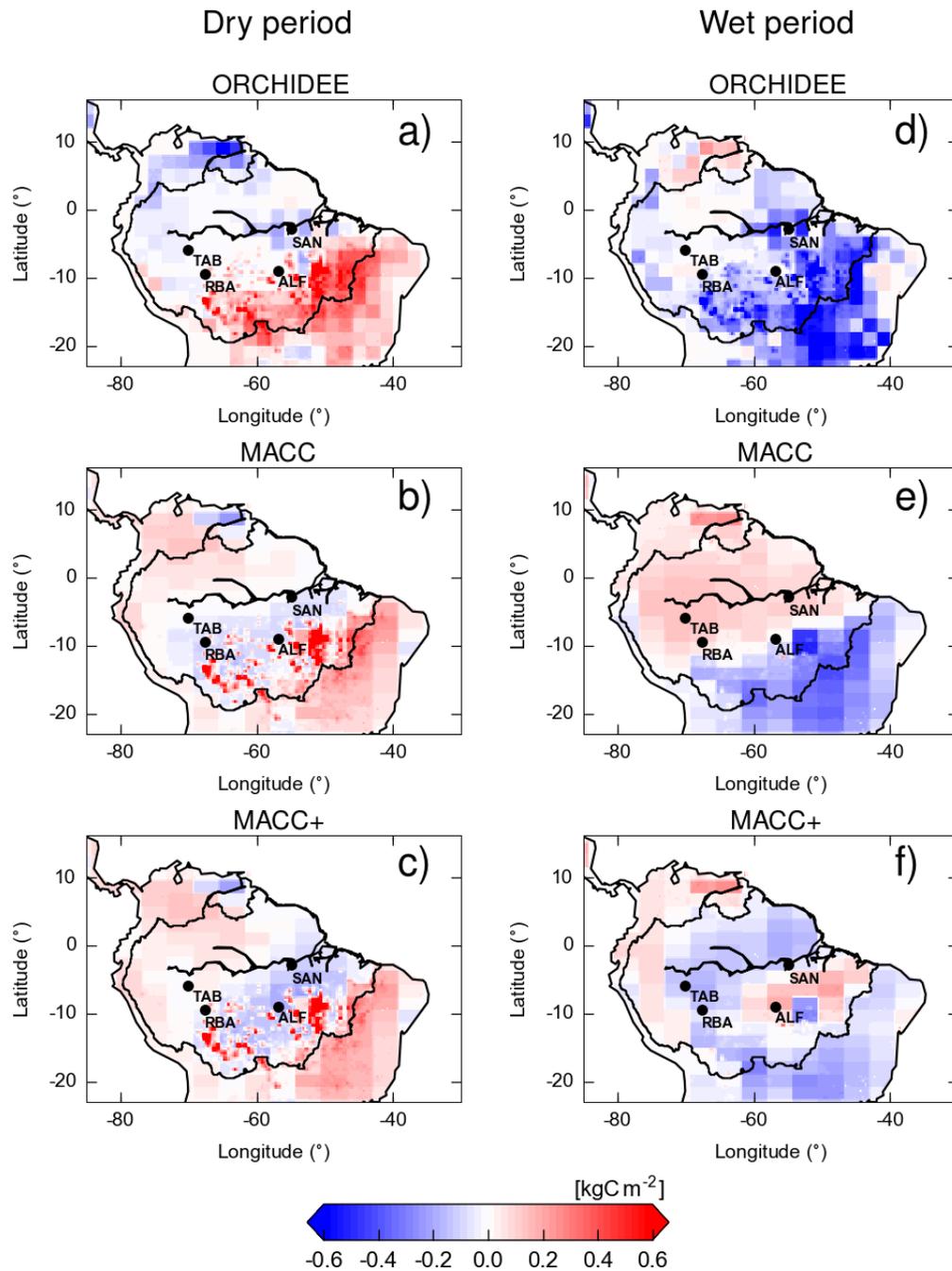


Figure 3.2: Mean seasonal distribution of total CO<sub>2</sub> fluxes over tropical South America land surface for 2010. Total flux is the sum of NEE and biomass burning and anthropogenic CO<sub>2</sub> emissions. Three estimates were built by combining NEE from ORCHIDEE (a, d), MACC (b, e) and MACC+ (c, f) with biomass burning emissions from GFED3.1 and fossil fuel emissions from EDGAR 4.2 FT database. Fluxes were averaged during the dry (July – October) (left column) and wet periods (November – June) (right column). Black dots: observation sites of GA2014.

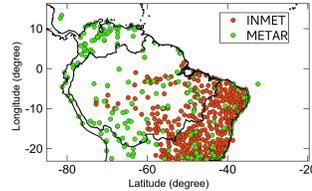


Figure 3.3: Location of sites with radiosonde data for model validation during 2010. Source: <http://esrl.noaa.gov/raobs>

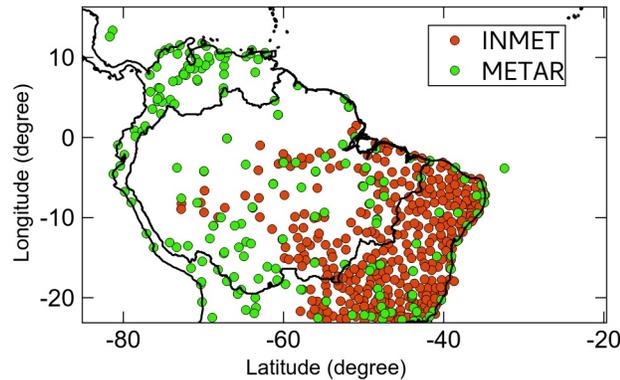


Figure 3.4: Location of surface meteorological stations used for model evaluation during 2010. Red: meteorological stations administered by the Brazilian National Institute of Meteorology (INMET). Green: meteorological stations installed at airports (METAR).

ally located at airports and discrete observations are made typically at 00:00 and 12:00 UTC. Following [Bela et al. \(2015\)](#), BRAMS and ECMWF were evaluated during two different meteorological regimes: the wet-to-dry (April – May) and the dry-to-wet (October – November) transitions seasons in 2010. At each site the root mean square error (RMSE) between observed and simulated time series was calculated for vertical profiles of potential temperature, dew point temperature, wind speed and wind direction at 7 vertical pressure levels between the surface and 300 hPa.

BRAMS and ECMWF were also evaluated against surface, regional meteorological station data from 152 meteorological stations installed at airports (METAR network) and 311 stations administered by the National Institute of Meteorology (INMET) of Brazil (Figure 3.4). Simulated fields of air temperature and air dew point temperature at 2 m, wind speed at 10 m and precipitation were interpolated to the station locations and the corresponding time series were compared to the measurements. Statistics of the misfits between simulations and observations were calculated following [Moreira et al. \(2013\)](#).

Simulated precipitation was compared to satellite data from the Tropical Rainfall Measuring Mission (TRMM) product 3b43v7, namely monthly mean precipitation rates (in  $\text{mm hr}^{-1}$ ) at  $0.25^\circ$  resolution. This product merges information from the three rainfall measuring instruments onboard the TRMM platform, the Precipitation Radar (PR), TRMM

Microwave Image (TMI), and Visible and Infrared Scanner (VIRS), with monthly totals from the Global Precipitation Climatology Centre (GPCC) rain gauge analysis. The more complete spatial and temporal coverage of this TRMM product allows to evaluate modeled spatio-temporal rainfall patterns better than with discrete surface stations.

### 3.1.3 Spatial and temporal CO<sub>2</sub> distribution and comparison to the regular vertical profiles from GA2014

In order to assess whether the transport models simulated the CO<sub>2</sub> distribution realistically, I evaluated the CO<sub>2</sub> transported with BRAMS and ECMWF with data from the aircraft profiles of GA2014. These data consisted of bi-weekly vertical profiles of CO<sub>2</sub> at four sites within the Amazon Basin (Figure 3.5): Alta Floresta (ALF), Rio Branco (RBA), Santarém (SAN) and Tabatinga (TAB).

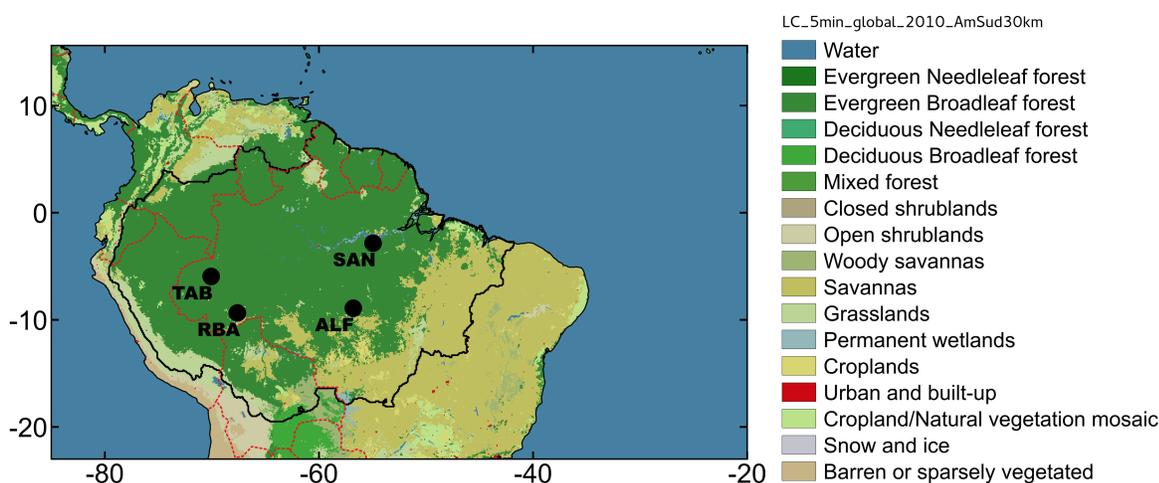


Figure 3.5: Aircraft sampling sites reported by Gatti et al. (2014) (black circles) and land cover type in tropical South America (Friedl et al., 2010).

The CO<sub>2</sub> measurements consisted of flask air samples taken in a descending spiral from ~4.4 km to approximately 300 m above sea level (a.s.l.) (close to the forest canopy), typically between 12:00 and 13:00 local time (LT). The simulated fields were sampled by selecting the model grid-cells corresponding to the horizontal and vertical positions of the observations, at the time when the measurements were taken. I compared simulated and observed mean vertical profiles for two contrasting periods, in terms of precipitation, defined after GA2014: the dry (July – October) and wet (November – June) seasons of 2010. In Amazonia, vertical tracer transport is fairly influenced by deep convection. To evaluate the sensitivity of the CO<sub>2</sub> vertical profiles to deep convection parameterization, transport was simulated using both meteorological models with the same set of fluxes and initial and boundary conditions, but suppressing deep convection in CHIMERE.

In addition to CO<sub>2</sub> at individual profile sites, I also compared model and observed CO<sub>2</sub> horizontal gradients between pairs of vertical profile sites along the dominant wind direction. CO<sub>2</sub> gradients between sites should be less sensitive to external fluxes and more influenced by the NEE in the areas between the two sites (Bréon et al., 2015; Boon et al., 2016; Staufer et al., 2016), and in view to increase the robustness of the inversions, I as-

simulated these gradients instead of individual data at each profile in the following chapter. Horizontal gradients were estimated along the three transects: TAB – SAN, RBA – ALF and RBA – SAN (Figure 3.6). The transects were chosen based on the dominant flow direction in the low- to mid-troposphere across the Amazon, with air entering from the equatorial Atlantic Ocean and flowing over the basin. Due to differences in the times and altitudes where observations were collected, data from each site were grouped into five altitude layers (or bins) every 1000 m, from the surface up to 5 km. Then, mean seasonal gradients were calculated as the mean CO<sub>2</sub> difference between two sites for each bin during the dry and wet seasons.

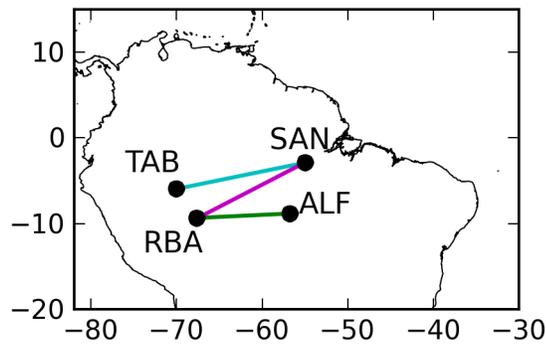


Figure 3.6: Horizontal transects established to evaluate horizontal CO<sub>2</sub> gradients.

I also analyzed two additional types of CO<sub>2</sub> gradients simulated with both BRAMS and ECMWF transport fields, with different NEE products: the horizontal gradient from 0 to 1 km a.s.l. (assuming that 1 km is the typical height of the planetary boundary layer (PBL) in the Amazon; (Fisch et al., 2004)), and the integral of the horizontal gradients from 0 to 5 km. These other types of horizontal gradients should be less sensitive to the vertical transport modeling errors than individual data, and than the horizontal gradients for different vertical layer. Inversions assimilating these types of gradients were also conducted in the following chapter.

## 3.2 Results

### 3.2.1 Simulated vs. observed meteorology

#### Simulated meteorology vs. vertical profiles from radiosondes

I calculated the RMSE of the misfits between the observed and simulated time series of potential and dew point temperature, wind speed and wind direction, at 7 levels between the surface and 300 hPa, for the 11 radiosonde observation sites shown in Figure 3.3, for both BRAMS and ECMWF. I summarized the results of the 11 sites with a box-and-whiskers plot for the wet-to-dry (April – May) (Figure 3.7) and dry-to-wet (October – November) (Figure 3.8) transition seasons of 2010. The plots characterize the distribution of the RMSE across the sites through the first and third quartiles (horizontal bars), the median value (thick vertical lines within the horizontal bars), the minimum and maximum values (whiskers) and outliers (black dots).

During April – May, the wet-to-dry transition, median RMSE in potential temperature simulated with both atmospheric models (Figure 3.7a) was largest at the surface ( $\sim 1.6$  K), and decreased to  $\sim 1$  K in the upper levels. For wind direction, RMSE was very large near the surface (up to  $100^\circ$  deviation), and decreased to  $\sim 50 - 100^\circ$  in altitude with both models (Figure 3.7d). Errors in dew point temperature and wind speed near the surface ( $\sim 1.6 - 2$  K and  $1 - 1.8$  m s<sup>-1</sup>, respectively) were lower than in upper levels ( $\sim 9 - 12$  K and  $2.5 - 4$  m s<sup>-1</sup>, respectively). During October – November (dry-to-wet transition) RMSE decreased with altitude for potential temperature (Figure 3.8a) and wind direction (Figure 3.8d), but increased for wind speed and dew point temperature. Overall, misfits between observations and ECMWF simulations showed higher median RMSE, and exhibited more variability (spread between 25th and 75th percentiles) in the upper levels of the atmosphere than for BRAMS, except for wind direction. Misfits between observations and model simulations with BRAMS showed more variability (spread between 25th and 75th percentiles) at the surface, except for dew point temperature in both transition seasons, and in wind direction in the dry-to-wet transition.

#### Simulated meteorology vs. ground-based data

Table 3.3 summarizes the RMSE and standard deviation of the misfits ( $STD_{MISFIT}$ ) between modeled and observations of temperature at 2 m (T), dew point temperature at 2 m ( $T_d$ ), precipitation at the surface and wind speed at 10 m ( $w_{10m}$ ). The difference between mean RMSE,  $STD_{MISFIT}$  and bias for each variable indicates that a large contribution to RMSE stemmed from the models' deficiency to reproduce the variability in the observations. In addition, near the surface BRAMS was drier (the bias in  $T_d$  was negative) during the wet-to-dry transition, and more humid in the dry-to-wet (the bias in  $T_d$  was positive), than ECMWF. Surface was warmer (bias of T > 0) in BRAMS than ECMWF during both periods. Both models overestimated the observed precipitation (bias > 0), with ECMWF predicting higher precipitation than BRAMS. At the surface, the RMSE and  $STD_{MISFIT}$  of wind speed—a critical variable for the transport of CO<sub>2</sub>—were similar ( $\sim 1.6$  and  $1.0$  m s<sup>-1</sup> respectively) for both models, independent of the transition period. BRAMS underestimated wind speed (bias < 0) and ECMWF overestimated it. Overall the different statistics in Table 3.3 show similar performance in reproducing the observed meteorology. In the discussion we will see that the models' performance is similar to that achieved in other studies of validation of meteorological models.

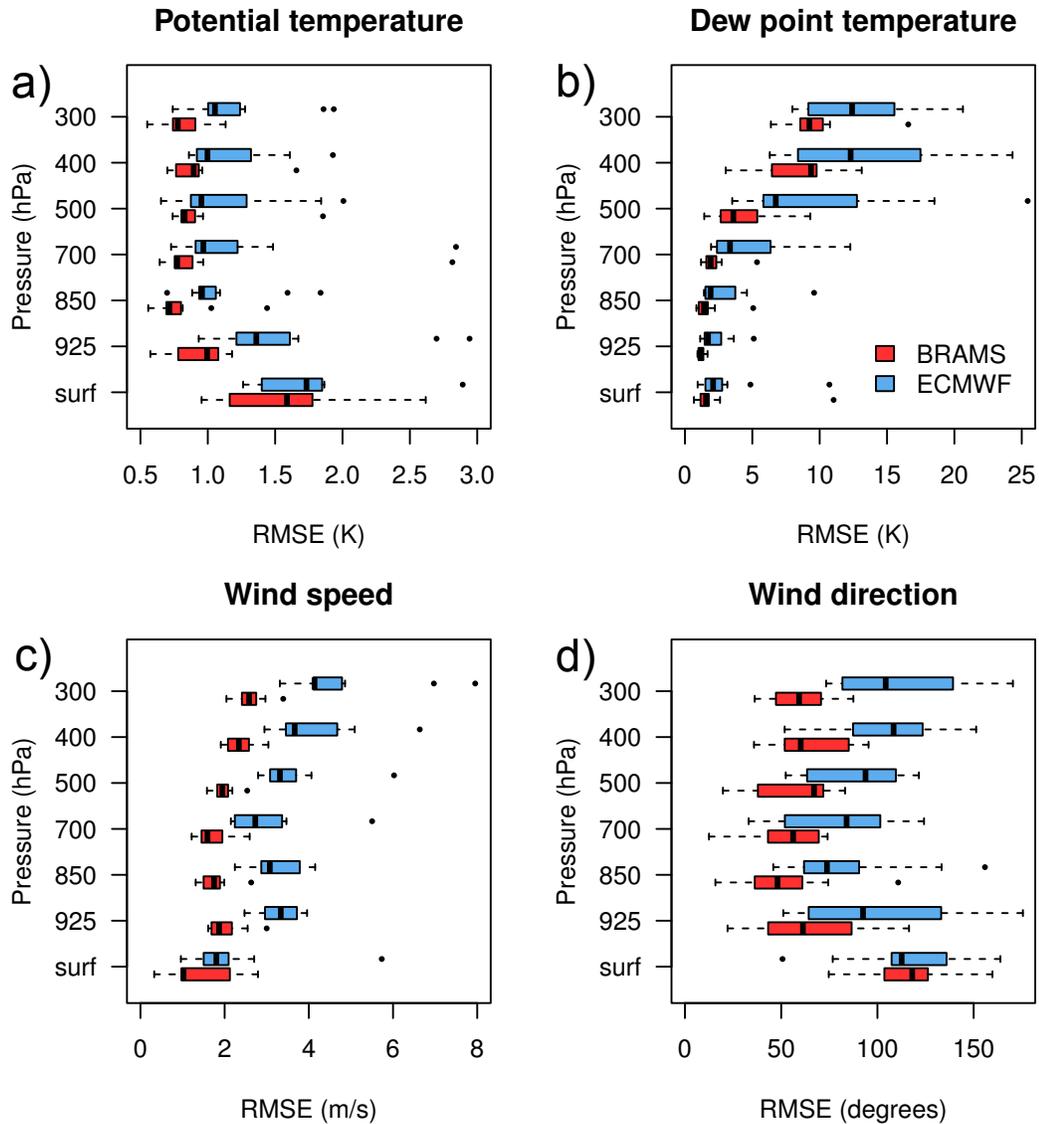


Figure 3.7: Distribution of the root-mean-square error (RMSE) between observations and simulations of (a) air temperature, (b) wind speed, (c) dew point temperature and (d) wind direction at 11 radiosonde observation sites (Figure 3.3) during April – May 2010 (wet-to-dry transition season). The observations correspond to discrete measurements at 00:00 or 12:00 UTC. The horizontal bars represent the distribution of the RMSE between the first and third quartiles. Thick black lines within the boxes represent the median RMSE. The whiskers indicate the minimum and maximum RMSE. Black dots represent the outliers.

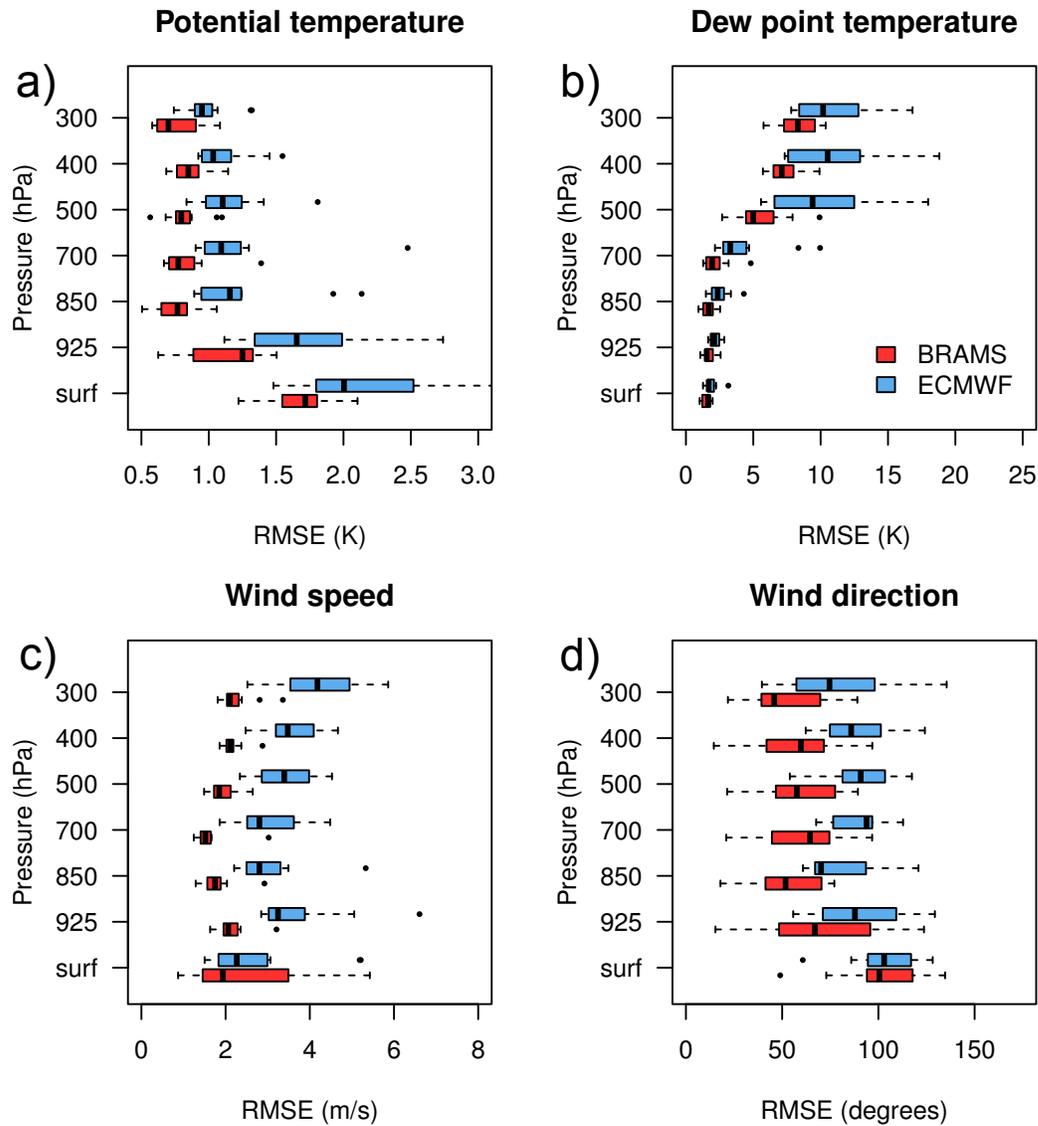


Figure 3.8: Distribution of the root-mean-square error (RMSE) between observations and simulations of (a) air temperature, (b) wind speed, (c) dew point temperature and (d) wind direction at 11 radisonde observation sites (Figure 3.3) during October – November 2010 (dry-to-wet transition season). The observations correspond to discrete measurements at 00:00 or 12:00 UTC. The horizontal bars represent the distribution of the RMSE between the first and third quartiles. Thick black lines within the boxes represent the median RMSE. The whiskers indicate the minimum and maximum RMSE. Black dots represent the outliers.

Table 3.3: Statistics of the misfits between observed and simulated surface air temperature (T), dew point air temperature ( $T_d$ ), wind speed and precipitation. Model simulations were compared to data from ground stations of the meteorological network in South America. Mean root-mean-square error (RMSE) and bias, and standard deviation of the model-observation misfits ( $STD_{MISFIT}$ ) were calculated for the periods April – May and October – November 2010.

		BRAMS		ECMWF	
		Apr-May	Oct-Nov	Apr-May	Oct-Nov
$T_d$ (K)	RMSE	1.80	1.88	1.47	1.69
	$STD_{MISFIT}$	1.19	1.21	1.07	1.17
	BIAS	-0.18	0.66	-0.13	0.52
T (K)	RMSE	2.05	2.03	1.79	1.72
	$STD_{MISFIT}$	1.23	1.21	1.18	1.19
	BIAS	0.81	0.70	-0.52	-0.26
Precipitation (mm)	RMSE	1.96	-	1.93	-
	$STD_{MISFIT}$	1.56	-	1.54	-
	BIAS	0.24	-	0.61	-
$w_{10m}$ (m/s)	RMSE	1.53	1.58	1.58	1.60
	$STD_{MISFIT}$	0.97	0.95	1.06	1.04
	BIAS	-0.18	-0.43	0.75	0.77

### Simulated precipitation vs. remote sensing data

Figure 3.9 compares monthly mean precipitation from TRMM and from both models during January and July 2010. Two patterns can be identified in TRMM. There is a NW-SE precipitation band over most of South America during January. This pattern corresponds to the South American Convergence Zone (SACZ), causing a belt of precipitation particularly intense during the austral summer (January – March) usually associated with strong convection in central South America, and intense precipitation across the southern part of the continent (Garreaud et al., 2009). During the austral winter (July – September) there is an E-W precipitation band extending over the tropical Atlantic Ocean and northern South America. Over the Equatorial oceans this pattern corresponds to the Intertropical Convergence Zone (ITCZ), associated with the wet season in northeastern Brazil (Garreaud et al., 2009). Figure 3.9 shows that both atmospheric models reproduce these patterns fairly well. This gives some confidence in the ability of both models to correctly locate convection areas that influence the vertical transport of  $CO_2$ . I calculated the monthly mean precipitation rate for three zones in tropical South America (see Figure 3.9): northern Amazon (region 1), equatorial Atlantic Ocean (region 2), and southern Amazon (region 3). The results, summarized in Table 3.4 shows that in January southern Amazon (zone 3) was the most humid region, according to TRMM, followed by northern Amazon (region 1) and the Equatorial Atlantic (region 3). ECMWF reproduced this pattern, but not BRAMS, which showed lower precipitation in zone 3. Compared to TRMM, precipitation over the simulation domain was  $\sim 18 - 23\%$  lower with ECMWF, and  $\sim 25 - 59\%$  lower with BRAMS. In July, precipitation in TRMM was high over northern Amazon and the Atlantic, whereas the southern Amazon showed low precipitations typical of the dry season in this region. Both models reproduced this pattern, but compared to TRMM, precipitation was  $\sim 23 - 31\%$  lower in ECMWF and  $\sim 49 - 65\%$  lower in BRAMS. Overall, both models reproduced well the seasonal precipitation distribution, except for BRAMS in January. On the other hand,

both models underestimated rainfall, with BRAMS showing lower precipitation.

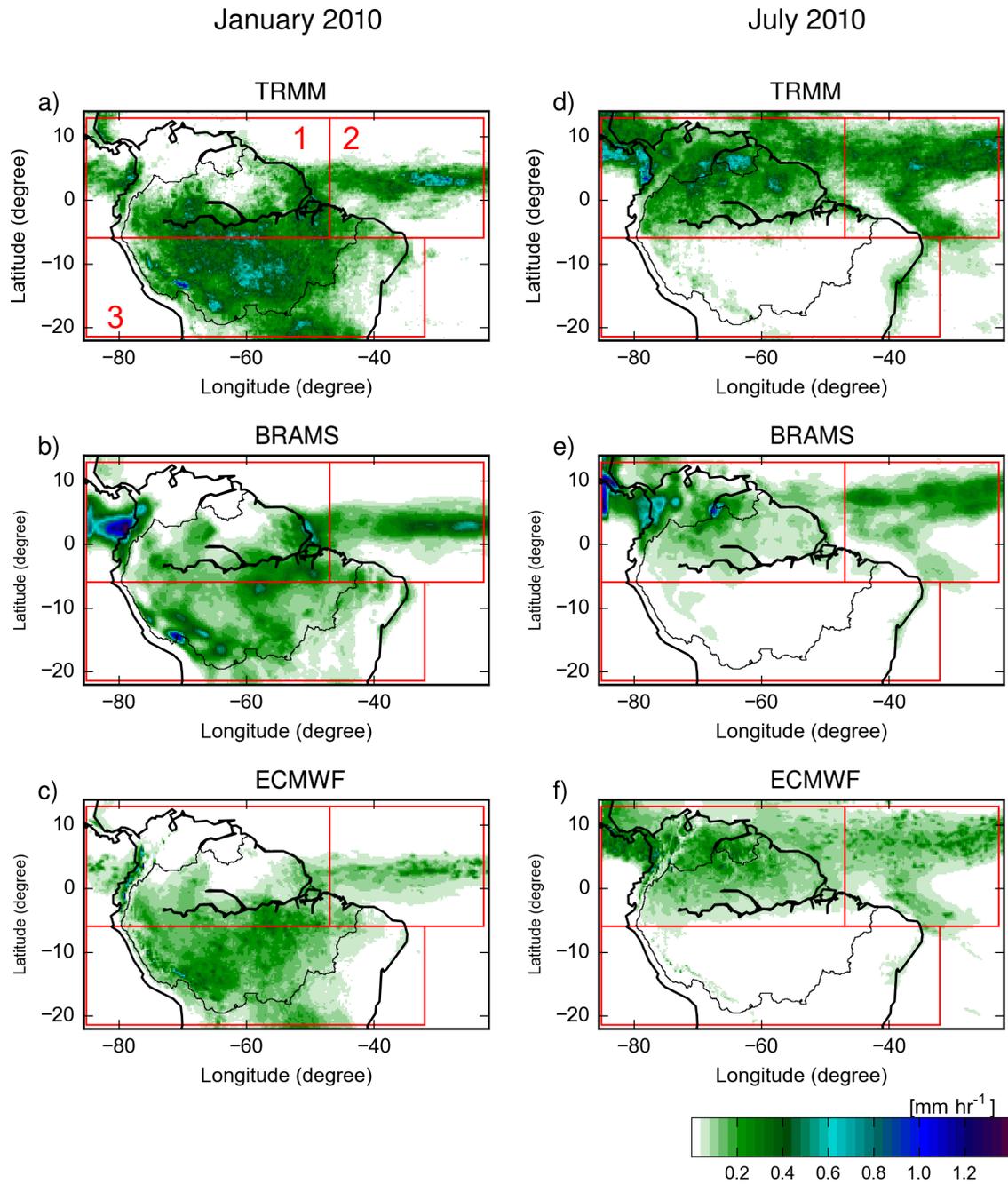


Figure 3.9: Monthly mean precipitation rate (in  $\text{mm hr}^{-1}$ ) during (a – c) January and (d – f) July. Precipitation rate has been calculated for three regions depicted in red boxes: (1) northern Amazon, (2) equatorial Atlantic Ocean and (3) southern Amazon. See Table 3.4.

Table 3.4: Monthly mean precipitation ( $\text{mm hr}^{-1}$ ) rate for TRMM, BRAMS and ECMWF for the three regions depicted in Figure 3.9: (1) northern Amazon, (2) equatorial Atlantic Ocean, (3) southern Amazon.

	January 2010			July 2010		
	TRMM	BRAMS	ECMWF	TRMM	BRAMS	ECMWF
Northern Amazon	0.146	0.11	0.119	0.315	0.16	0.241
Equatorial Atlantic	0.108	0.076	0.083	0.245	0.119	0.169
Southern Amazon	0.227	0.092	0.184	0.031	0.011	0.02

### 3.2.2 Spatial and temporal $\text{CO}_2$ distribution

#### Distribution of $\text{CO}_2$ along vertical cross-Amazon transects

Figure 3.10 shows the mean seasonal averaged  $\text{CO}_2$  column ( $\text{XCO}_2$ ), obtained from afternoon values (12:00 – 18:00 LT), for the dry (July – October) and wet (November – June) seasons of 2010. Transport was simulated with both CH-BRAMS and CH-ECMWF, using the ORCHIDEE NEE, fire  $\text{CO}_2$  emissions GFED3.1, fossil fuel  $\text{CO}_2$  emissions EDGAR4.2, ocean fluxes TAKA09 and BC\_MACC+ boundary conditions. 12:00 – 18:00 LT is the time window used for the selection of the  $\text{CO}_2$  data to be assimilated in the next chapter. I made this selection because modeling high nighttime concentrations close to the surface is problematic (see Chapter 2). The accumulation of  $\text{CO}_2$  in the nocturnal shallow boundary layer strongly depends on the local fluxes and stability of the PBL and mesoscale models, as well as global transport models, have difficulties to resolve these local effects (Karstens et al., 2006).

During the dry period, a positive N-S  $\text{CO}_2$  gradient was modeled over the Amazon with both CH-BRAMS and CH-ECMWF. High concentrations were simulated along the southeastern basin in the areas covered by savanna vegetation and savanna-forest transition areas (Figure 3.5). These areas exhibit intense fire emissions included in GFED3.1. These areas are typically subject to intense fire activity due to land cover conversion and to other agricultural practices (Pivello, 2011). During the wet period, both models predicted a negative NW-SE gradient, with low concentrations over the southern Amazon basin, especially over the areas covered with savanna vegetation, where the NEE from ORCHIDEE showed intense  $\text{CO}_2$  uptake (Figure 3.2). Note that, in general, during both the dry and wet periods, CH-BRAMS predicts lower  $\text{CO}_2$  across the continental land than CH-ECMWF. Given the overall good agreement between mean horizontal wind fields for both models during both dry and wet periods (Figure 3.11) the differences between  $\text{XCO}_2$  fields predicted with CH-BRAMS and CH-ECMWF reflect differences in the strength of the vertical transport coupled with differences in advection (strength and/or direction) between the models.

Figures Figure 3.12 and Figure 3.13 show vertical  $\text{CO}_2$  cross-sections along two transects drawn on Figure 3.10, for both dry and wet periods, respectively. Transect 1 runs NW-SE from  $7^\circ$  N,  $80^\circ$  W to  $20^\circ$  S,  $40^\circ$  W. Transect 2 runs along  $7^\circ$  S from  $80^\circ$  W to  $40^\circ$  W.

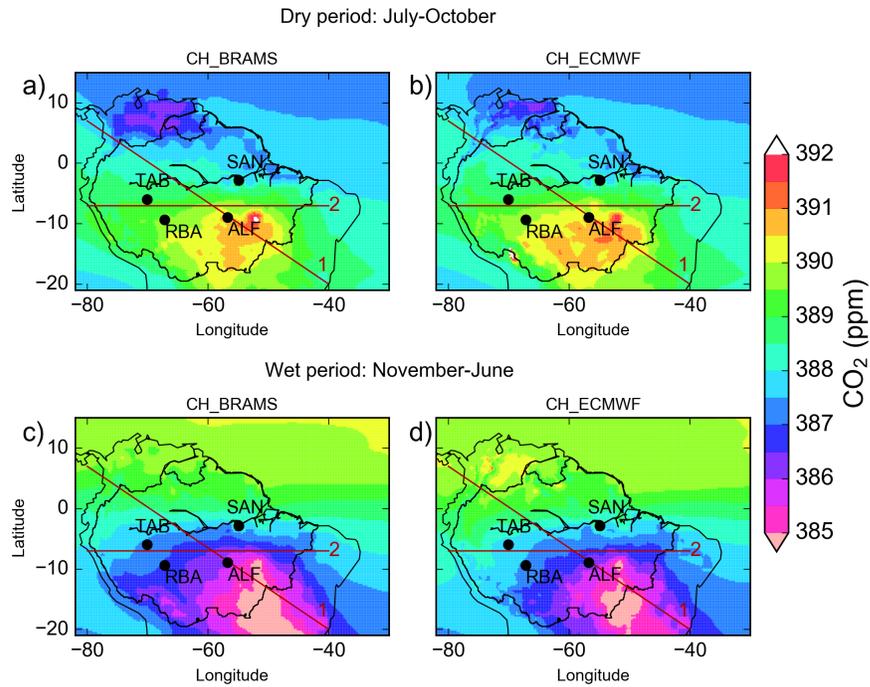


Figure 3.10: Mean seasonal CO<sub>2</sub> column during the (a, b) dry (July – October) and (c, d) wet (November – June) periods of 2010, simulated with CH-BRAMS and CH-ECMWF. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

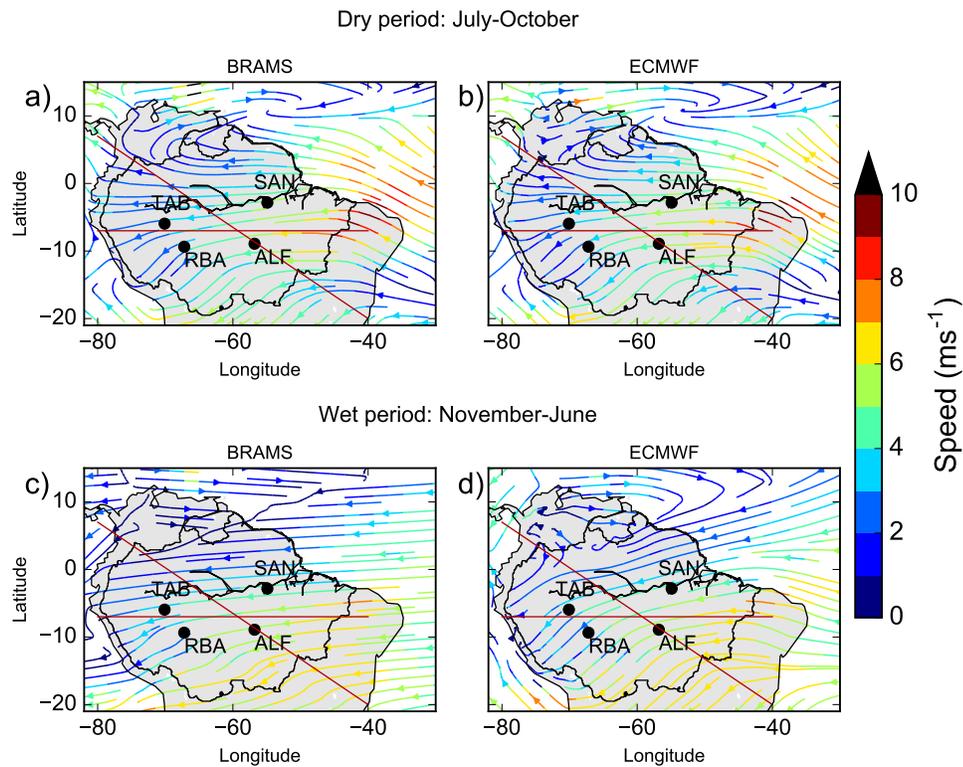


Figure 3.11: Mean seasonal wind fields during the (a, b) dry (July – October) and (c, d) wet (November – June) periods of 2010, simulated with BRAMS and ECMWF. Wind fields averaged between the surface and the top of the transport model (~10 km).

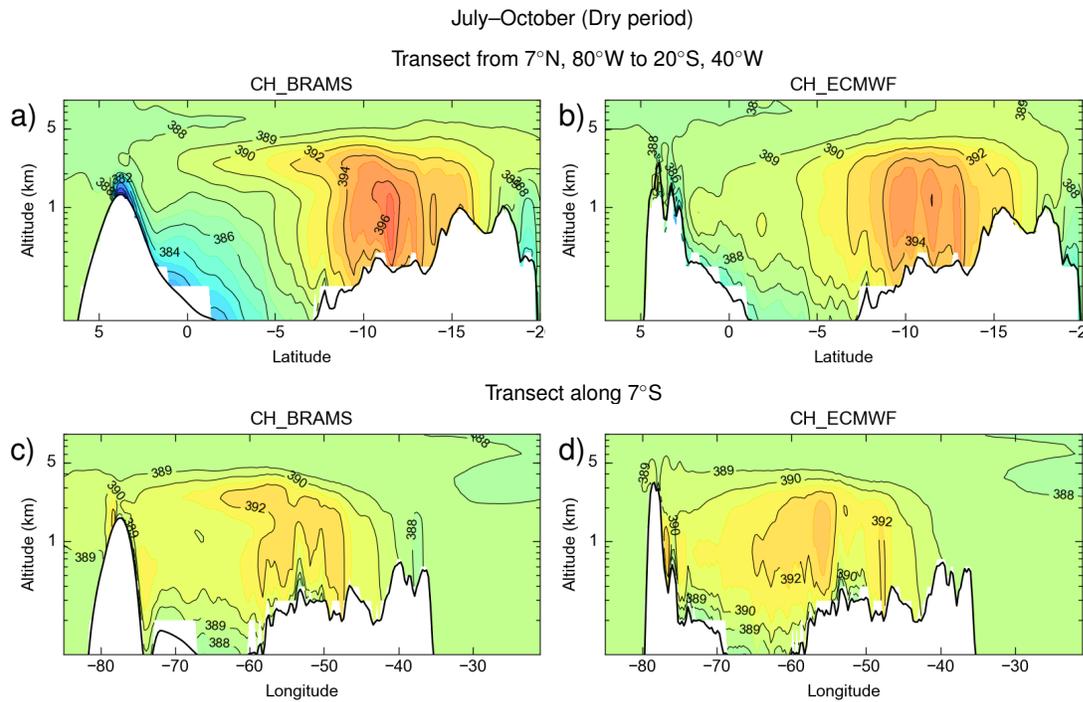


Figure 3.12: Vertical cross-sections of the mean simulated CO<sub>2</sub> (in ppm) for July – October 2010 (dry period) for CH-BRAMS and CH-ECMWF along (a, b) the transect from 7° N, 80° W to 20° S, 40° W, and along (c, d) 7° S. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

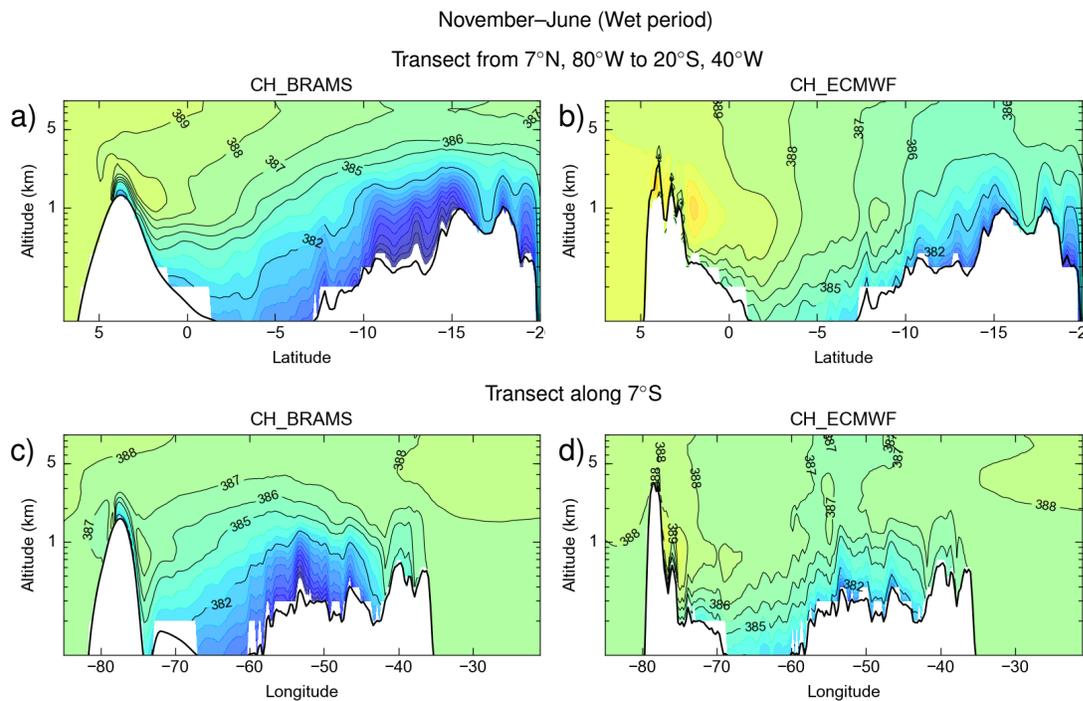


Figure 3.13: Vertical cross-sections of the mean simulated CO<sub>2</sub> (in ppm) for November–June 2010 (wet period) for CH-BRAMS and CH-ECMWF along (a, b) the transect from 7° N, 80° W to 20° S, 40° W, and along (c, d) along 7° S. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

During the dry period along Transect 1 (Figure 3.12), in particular between  $\sim 3^\circ$  N and  $3^\circ$  S, air with lower CO<sub>2</sub> reached higher altitudes in CH-BRAMS than CH-ECMWF. Given that this part of the transect was located over a moderate CO<sub>2</sub> sink, this structure likely reflects a more intense vertical transport of CO<sub>2</sub>-depleted air in CH-BRAMS than in CH-ECMWF. Likewise, along  $5^\circ$  S –  $5^\circ$  S, over an area of intense CO<sub>2</sub> emission (Figure 3.2a), air rich in CO<sub>2</sub> was also transported higher in CH-BRAMS than in CH-ECMWF. During the wet period (Figure 3.13), the lowest concentrations were located over areas of strong uptake in ORCHIDEE (Figure 3.2b), and CO<sub>2</sub>-depleted air was also transported to higher altitude in CH-BRAMS than in CH-ECMWF.

During both dry and wet seasons, above 1 km CO<sub>2</sub> was distributed horizontally with more structure in CH-BRAMS than CH-ECMWF. The prescribed topography, which is different in both models, seemed to play a role in this structure. Along Transect 1, terrain relief at about  $3^\circ$  N is  $\sim 1$  km in CH-BRAMS, but  $\sim 3$  km high in CH-ECMWF. Higher terrain relief in CH-ECMWF might have acted as a barrier to CO<sub>2</sub> advection. This was also observed along Transect 2 at  $\sim 87^\circ$  W. Thus, the difference in model topography likely adds to the differences in the strength of the vertical transport between both models, coupled to differences in advection between the models, thereby generating differences in the horizontal distribution between CH-BRAMS and CH-ECMWF.

## Comparisons to vertical profiles of CO<sub>2</sub> measurements

### Dry season

Figure 3.14 compares observed mean, seasonal vertical CO<sub>2</sub> profiles and model simulations using CH-BRAMS and CH-ECMWF at the four aircraft sampling sites of GA2014 (see Figure 3.5), using ORCHIDEE NEE, fossil fuel CO<sub>2</sub> emissions from EDGAR4.2, biomass burning CO<sub>2</sub> emissions from GFED3.1, air-sea CO<sub>2</sub> exchanges from TAKA09 and BC\_MACC+ boundary conditions. The models were sampled at the time of the observation, at the grid-cell of CHIMERE where the observations were located. Observations and model simulations were averaged over July – October (dry season) and June – November (wet season), after GA2014. Overall, both models exhibited similar performance in reproducing the shape of the observed mean vertical profiles. RMSE values varied between  $\sim 1$  – 2 ppm for both models. Modeled mean profiles overestimated the variability of the observations during both seasons, especially during the dry season. Figure 3.14 illustrates that concentrations simulated with CH-ECMWF were closer to the observations near the surface, e.g. at TAB and ALF, but in some instances CH-BRAMS was better between 1 and 2 km, e.g. at TAB and ALF.

### Wet season

Models also had a similar performance to reproduce the mean observed profiles during the wet season; RMSE varied between  $\sim 1.4$  – 3.6 and 1.2 – 2.8 ppm for CH-BRAMS and CH-ECMWF, respectively. Both models systematically predicted lower-than-observed CO<sub>2</sub> along the profiles at all sites, but CH-BRAMS predicted lower concentrations than CH-ECMWF by  $\sim 0.8$  – 4.8 ppm. A misfit of  $\sim 2$  ppm at altitudes around 4 km a.s.l was seen in both model simulations. This misfit in altitude suggests a bias from the boundary conditions rather than NEE. This is explored in the following. CH-BRAMS exhibited lower misfits to the observations, e.g. at TAB, RBA and ALF, but the misfits increased relative to CH-ECMWF in altitude. In some instances, like at SAN, both models failed to reproduce

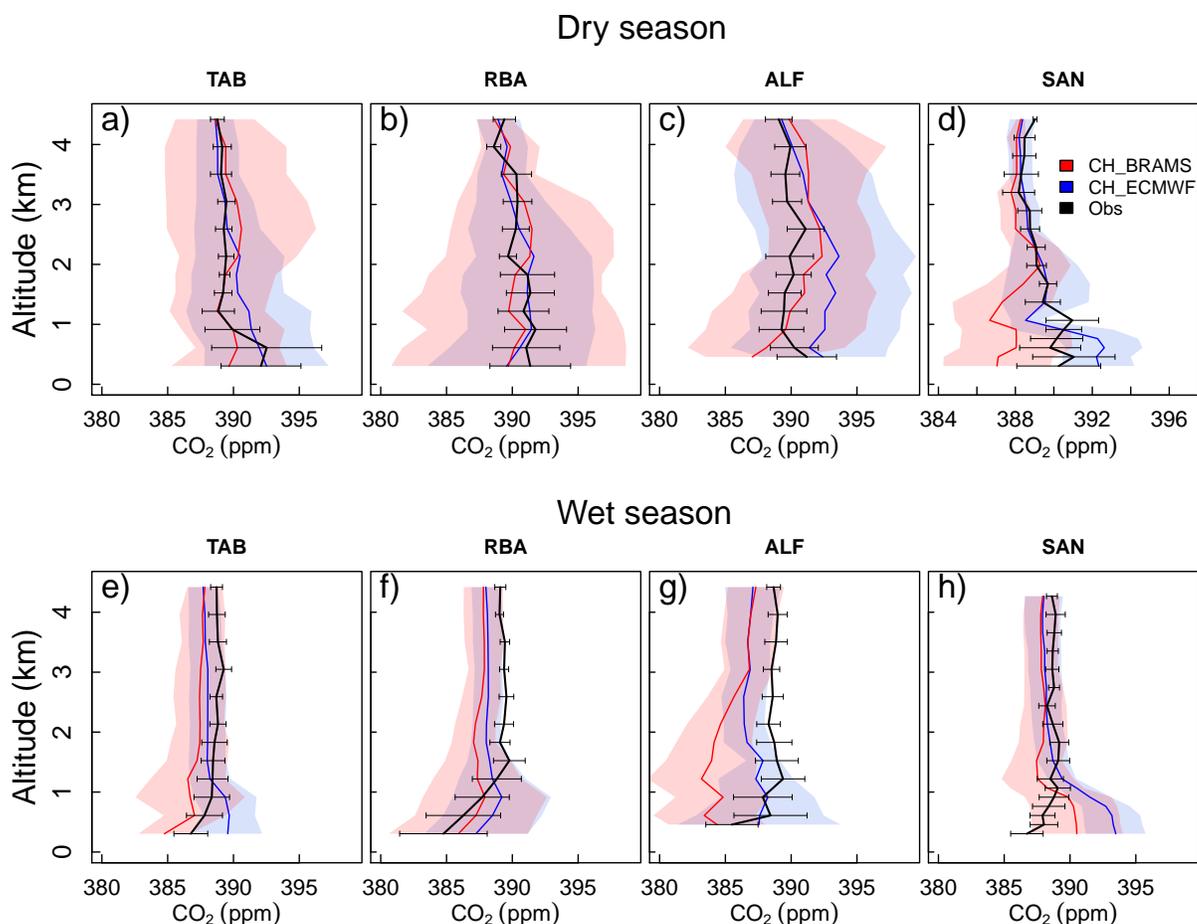


Figure 3.14: Observed (Black lines) and simulated mean CO<sub>2</sub> vertical profiles at the four sampling sites reported by [Gatti et al. \(2014\)](#) for 2010 during the dry (July – October) and wet (November – June) periods. CO<sub>2</sub> was simulated with CH-BRAMS (red lines) and CH-ECMWF (blue lines). Horizontal bars: spread of the observations ( $1-\sigma$ ). Shaded areas: spread of model simulation ( $1-\sigma$ ). Models were sampled at the time of the observation, and at the model grid cell where the observations were located. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

the observed CO<sub>2</sub> reduction toward the surface and predicted CO<sub>2</sub> buildup (Figure 3.14h). This behavior, though, could also be caused by errors in the simulation of surface fluxes by ORCHIDEE.

Although mean vertical profiles during both seasons showed an overall agreement to the observations with ORCHIDEE NEE, the models' ability to reproduce individual observed profiles was much lower—the models explained around 10% ( $R^2$ ) of the vertical and temporal variability of the observations at a given measurement site during a given season (Figure 3.15).

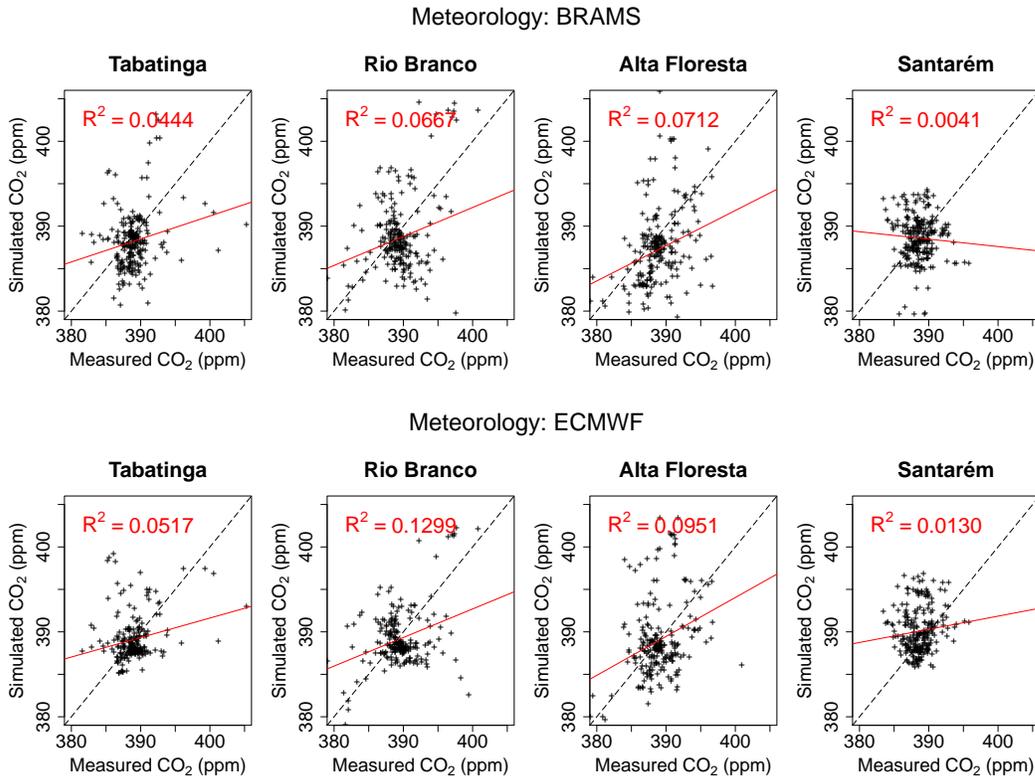


Figure 3.15: Model performance in reproducing the individual observations at each sampling site. No differentiation is made between dry and wet seasons because there are few points during the dry season. Number of observations available at each site: TAB = 221, RBA = 222, ALF = 222, SAN = 240. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

### BRAMS and ECMWF differences in vertical structure of CO<sub>2</sub> for transects between aircraft sites

Figures 3.12 and 3.13 show that differences in the strength of vertical transport between the models—coupled with differences in horizontal advection—caused CH-BRAMS to predict lower CO<sub>2</sub> along the vertical than CH-ECMWF during both dry and wet seasons. Vertical cross-sections between pairs of aircraft sites, namely TAB-SAN (TRAN<sub>TAB,SAN</sub>) (Figure 3.16) and RBA-ALF (TRAN<sub>RBA,ALF</sub>) (Figure 3.17), illustrate the influence of those mechanisms on CO<sub>2</sub> at the sampling sites.

During the dry period, along TRAN<sub>TAB,SAN</sub>, approximately between 75° W and 60° W, CO<sub>2</sub>-enriched air was transported higher in altitude in CH-BRAMS, and advected westward with respect to CH-ECMWF. Lower relief in CH-BRAMS, e.g. toward 80° W, suggests that CO<sub>2</sub> was more easily advected from the Amazon lowland across the Andes towards the Pacific ocean in comparison with CH-ECMWF, contributing to lower CO<sub>2</sub> in the profiles. Along TRAN<sub>RBA,ALF</sub> the vertical structure was similar between the two models east of ~57° W. To the west of ~57° W, however, CH-BRAMS transported CO<sub>2</sub> out of the continent more

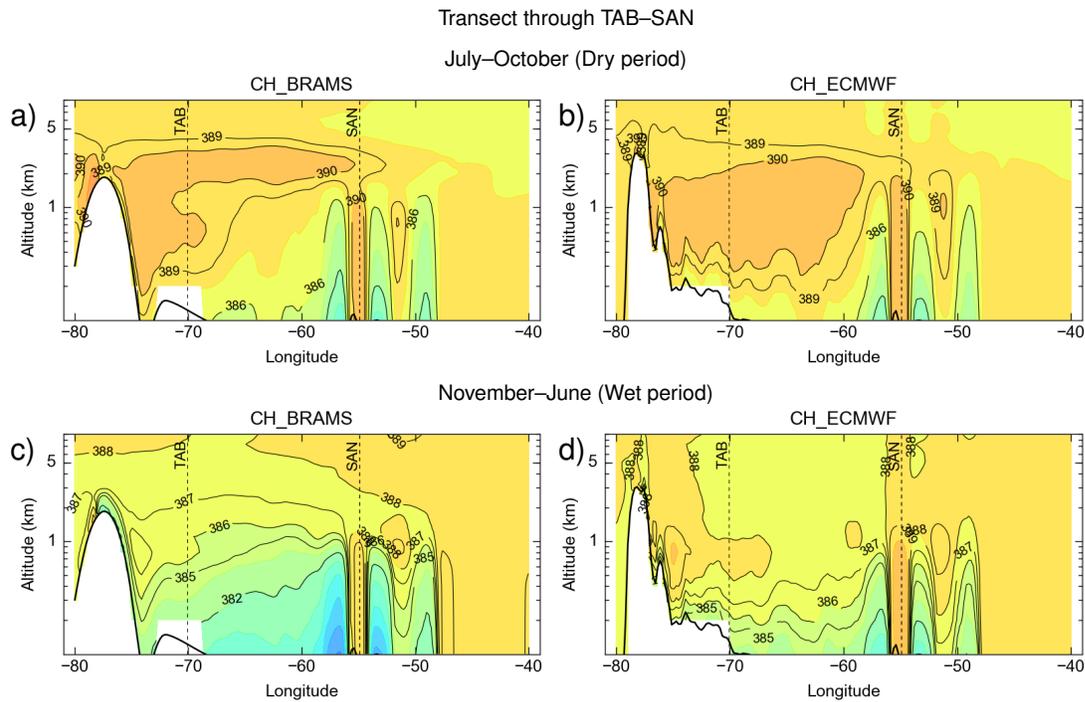


Figure 3.16: Vertical cross-section of the mean simulated CO<sub>2</sub> (in ppm) for July – October (dry period) and November – June (wet period) 2010 for CH-BRAMS and CH-ECMWF, along a transect through stations TAB and SAN (5.95° S, 70.1° W to 2.84° S, 54.9° W). Dashed vertical lines indicate longitudinal position of the stations. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

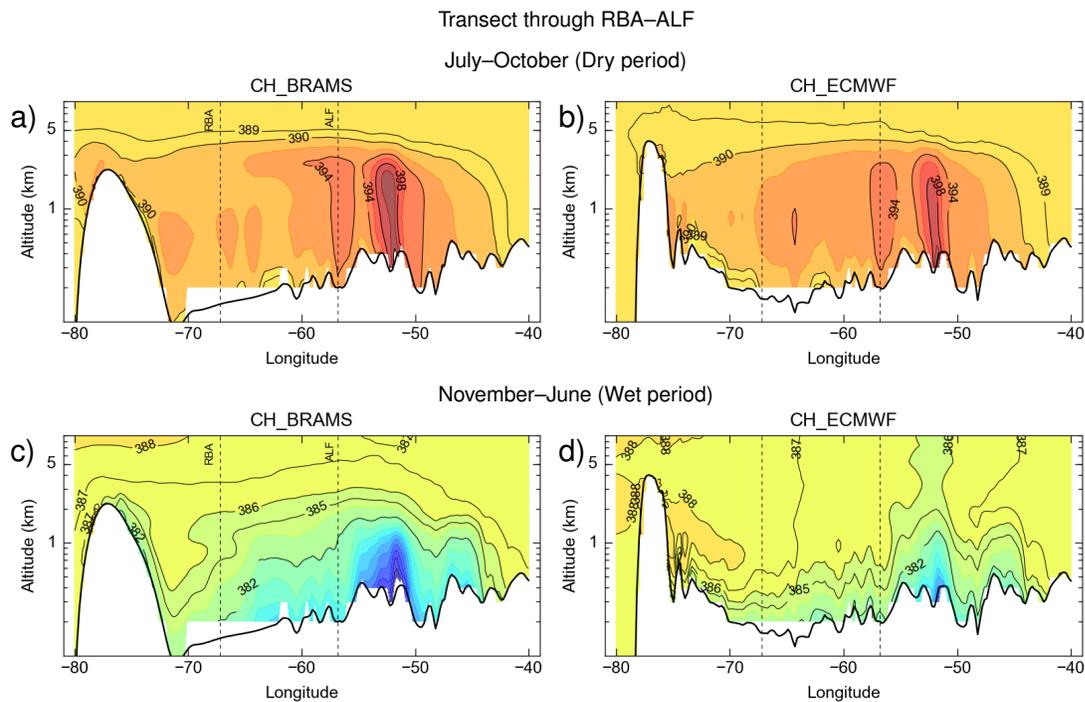


Figure 3.17: Vertical cross-section of the mean simulated CO<sub>2</sub> (in ppm) for July – October (dry period) and November – June (wet period) 2010 for CH-BRAMS and CH-ECMWF, along a transect through stations RBA and ALF (9.36° S, 67.2° W to 8.92° S, 56.8° W). Dashed vertical lines indicate longitudinal position of the stations. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

easily than ECMWF, likely due to lower terrain. Localized structures are observed about 60° W – 50° W along TRAN<sub>TAB,SAN</sub>, where air with lower CO<sub>2</sub> reached up to 1 km a.s.l. Those structures are related to stronger NEE uptake around SAN predicted by ORCHIDEE (Figure 3.2a) and illustrate the influence of local fluxes. Likewise, vertical transport of CO<sub>2</sub> was more intense in CH-BRAMS during the wet period. As a result, there was less CO<sub>2</sub> within the PBL than in CH-ECMWF. Lower terrain to the west of the transects in CH-BRAMS made it easier for large-scale advection to transport CO<sub>2</sub> out of the Amazon lowlands.

## Vertical CO<sub>2</sub> gradients between boundary layer and the free troposphere

### Dry season

The vertical CO<sub>2</sub> gradient between the boundary layer and the free troposphere (BL–FT) is an important indicator of vertical transport. I estimated the mean BL–FT difference from the simulated CO<sub>2</sub> difference between 300 m and 4 km a.s.l. for both observed and simulated profiles at each site (Table 3.5). Positive BL–FT differences were observed during the dry season, which reflects CO<sub>2</sub> accumulation in the BL and the surface being a source of CO<sub>2</sub> to the atmosphere. During the dry season, CH-BRAMS—using ORCHIDEE NEE—underestimated the observed positive BL–FT differences at all sites. CH-BRAMS predicted a positive difference at TAB and RBA, but a negative difference at ALF and SAN opposite to the observations. On the other hand, CH-ECMWF captured positive BL–FT differences in the dry season, but overestimated the magnitude of the gradients at all sites except RBA.

### Wet season

During the wet season, the observed differences were negative, indicative of CO<sub>2</sub> uptake. CH-BRAMS predicted negative BL–FT differences, in qualitative agreement with the observations, except at SAN, where the model predicted a positive BL–FT difference. Conversely, CH-ECMWF predicted a positive BL–FT difference, opposite to the observation, except at RBA during the wet season. Altogether, both models reproduced approximately the seasonality of the BL–FT gradient, except during the wet season for CH-ECMWF. Observed and modeled gradients, however, differed by 0.2 – 7 ppm. This result shows that even when prescribed with the same NEE, two different transport models can produce BL–FT differences of different sign, indicating large differences in BL mixing and exchange with the FT. At this stage, it is difficult to connect the unrealistic BL–FT differences during the wet season when using CH-ECMWF to this model only, since they can also be associated with errors in the NEE from ORCHIDEE. But since CH-BRAMS reproduced better both the magnitude and the direction of the observed BL–FT differences in the wet season, CH-ECMWF might be partly responsible for the misfits to the observations.

## Sensitivity of simulated CO<sub>2</sub> to convective transport

I studied CO<sub>2</sub> transport by deep convection by deactivating this process in the mass flux of CHIMERE. Figure 3.18 shows that without deep convection the simulated profiles can vary by ~1.6 ppm relative to the simulations with deep convection, with higher differences close to the surface during both seasons. Despite these differences, models did not predict substantial changes in the vertical CO<sub>2</sub> structure in comparison to observations in Figure 3.18, at least within the low- to mid-troposphere.

Table 3.5: Seasonal mean CO<sub>2</sub> difference between boundary layer (BL) and free troposphere (FT) at the four sampling sites in Amazonia.

Site	Dry season			Wet season		
	Obs.	CH-BRAMS	CH-ECMWF	Obs.	CH-BRAMS	CHECMWF
TAB	3.31	1.07	3.88	-1.96	-3.08	1.87
RBA	1.97	1.07	0.66	-4.32	-1.90	-0.75
ALF	2.13	-2.80	3.12	-3.17	-2.95	0.45
SAN	1.23	-1.24	3.97	-1.90	2.59	5.46

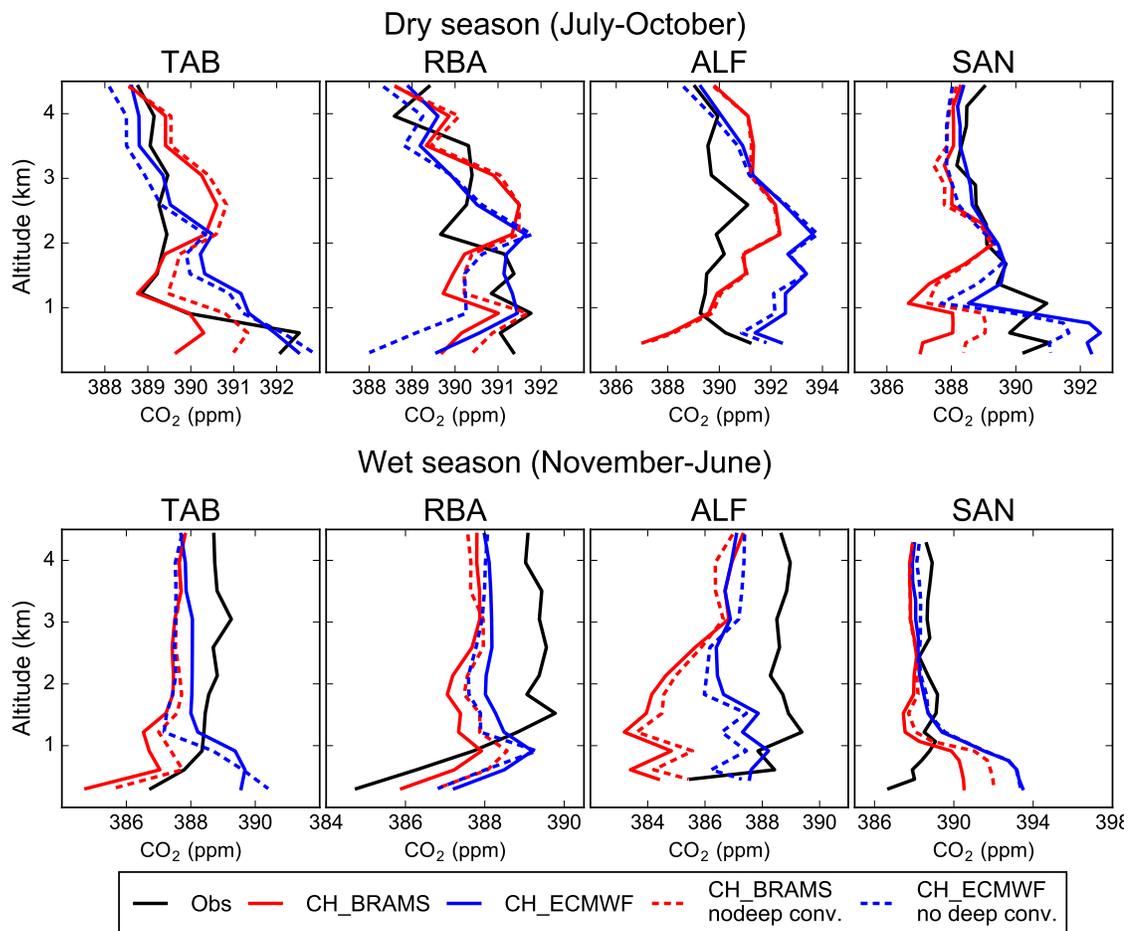


Figure 3.18: Sensitivity of simulated seasonal mean vertical CO<sub>2</sub> profiles to deep convection parameterization for July – October (dry season) and November – June (wet season). CO<sub>2</sub> was simulated with CH-BRAMS (red) and CH-ECMWF (blue), with the deep convection parameterization activated (solid lines) or suppressed (dashed lines). CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

### **Components of simulated CO<sub>2</sub> per type of fluxes in the simulation area, and their sensitivity to different lateral CO<sub>2</sub> boundary conditions and NEE fluxes inside the domain**

This section analyzes the individual impact of the different type of fluxes on the concentrations modeled at the aircraft measurement sites: that of NEE, CO<sub>2</sub> boundary conditions, biomass burning CO<sub>2</sub> emissions, ocean CO<sub>2</sub> fluxes and fossil fuel CO<sub>2</sub> emissions. Further, for NEE and CO<sub>2</sub> boundary conditions this impact was simulated using different products to model these fluxes or boundary conditions, in order to provide a typical characterization of the impact of uncertainties in the flux along with that of the impact of the fluxes. The sensitivity of the mean vertical CO<sub>2</sub> profiles at each site to different NEE and CO<sub>2</sub> boundary conditions is shown in Figures 3.19 to 3.22.

During both the dry or wet seasons, changes to NEE brought major impact on the vertical profiles, independent of the meteorological forcing used to simulate CO<sub>2</sub> transport. With the three sets of NEE products the largest signal from NEE took place within 0 – 2 km and reached up to 5 ppm. Above 2 km a.s.l., the signature of NEE did not exceed 1 ppm, except at ALF with CH-BRAMS during the wet season (Figure 3.20c), where the signature reached up to ~3 ppm. The difference between the signatures of any two sets of NEE ( $\Delta$ NEE)—a proxy for the uncertainty in the concentrations due to the uncertainty in the NEE—was larger below 2 km during both dry and wet periods. In the dry period,  $\Delta$ NEE reached up to ~5 ppm below 2 km, and  $\leq$  2 ppm above 2 km. There was the same behavior in the wet period, but  $\Delta$ NEE reached up to 9 ppm below 2 km with CH-BRAMS (Figure 3.20c). Differences between the signals of ORCHIDEE-MACC and ORCHIDEE-MACC+ reached up to 9 ppm, even though both sets of NEE were optimized in a global inversion. This illustrates the strong impact that the assimilation of the data of GA2014 had on NEE in the Amazon region. The second strongest influence on the concentrations was exerted by biomass burning emissions during the dry season (panels i – l in Figures 3.19 and Figure 3.21). The signal was strongest, in general, close to the surface, and values reached ~4 ppm. The exception is ALF, where concentrations were larger at ~2 km a.s.l. (Figures 3.19k and Figure 3.21k). Note that biomass burning emissions influenced the concentrations even in altitude. With both boundary conditions datasets, background concentrations were about 388 ppm across the sites, with little variation along the vertical, especially in the dry season (panels e – h in Figures 3.19 and 3.21). During the wet season, however, the signature of boundary conditions at sites like ALF and SAN showed more structure (panels g and h in Figures 3.20 and 3.22), and the signature varied by as much as 1 ppm between 4 km a.s.l. and the surface. The difference between the two sets of boundary conditions ( $\Delta$ BC) was lower than for NEE; it reached up to 1 ppm in the dry period and up to 0.5 ppm in the wet period. During both dry and periods, the signals of ocean fluxes and fossil fuel emission were almost zero.

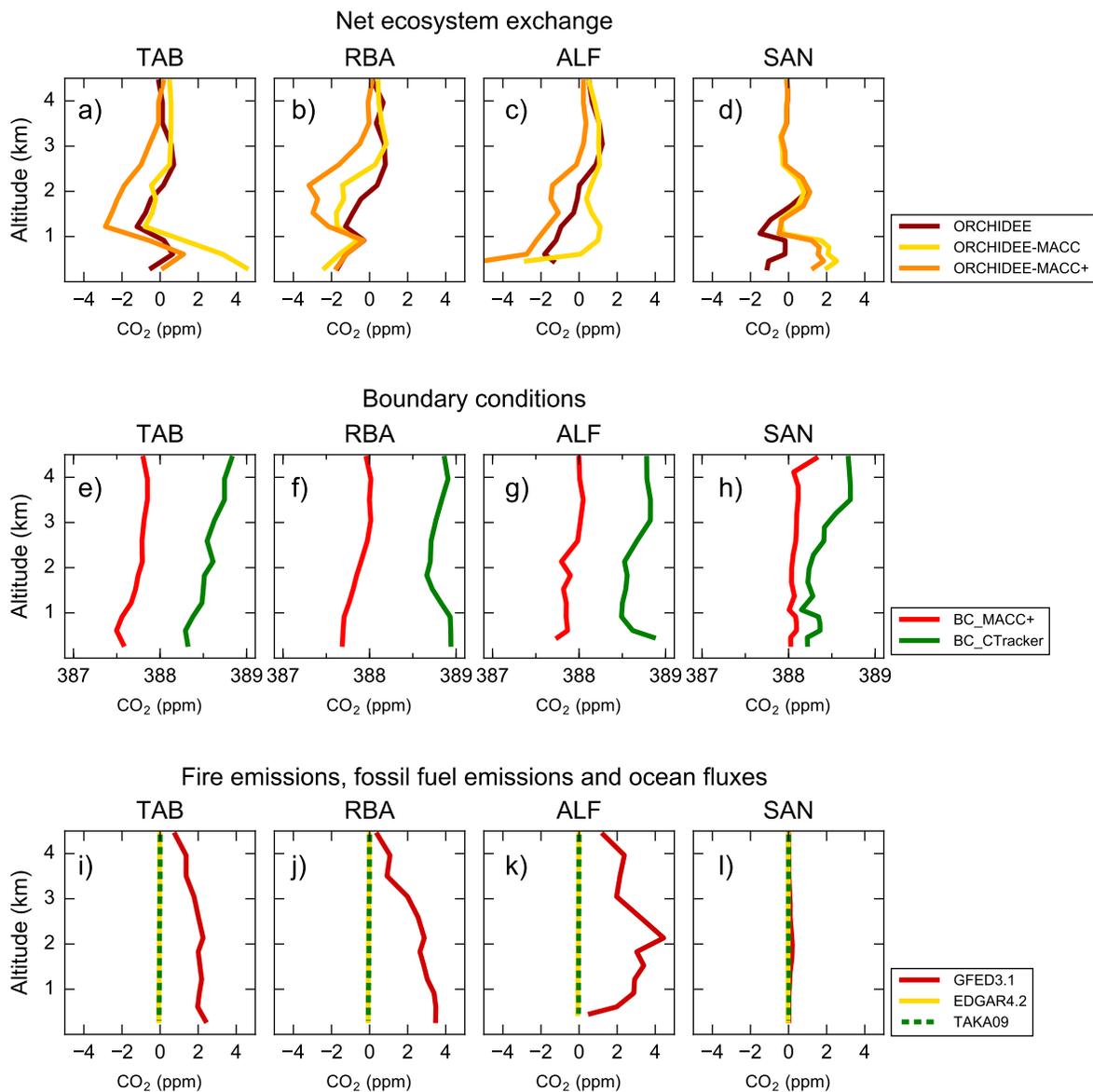


Figure 3.19: Sensitivity of seasonal mean vertical CO<sub>2</sub> profiles simulated with CH-BRAMS to (a – d) NEE, (e – h) boundary conditions and (e – h) fire and fossil fuel emissions and ocean fluxes for July – October (dry season) of 2010.

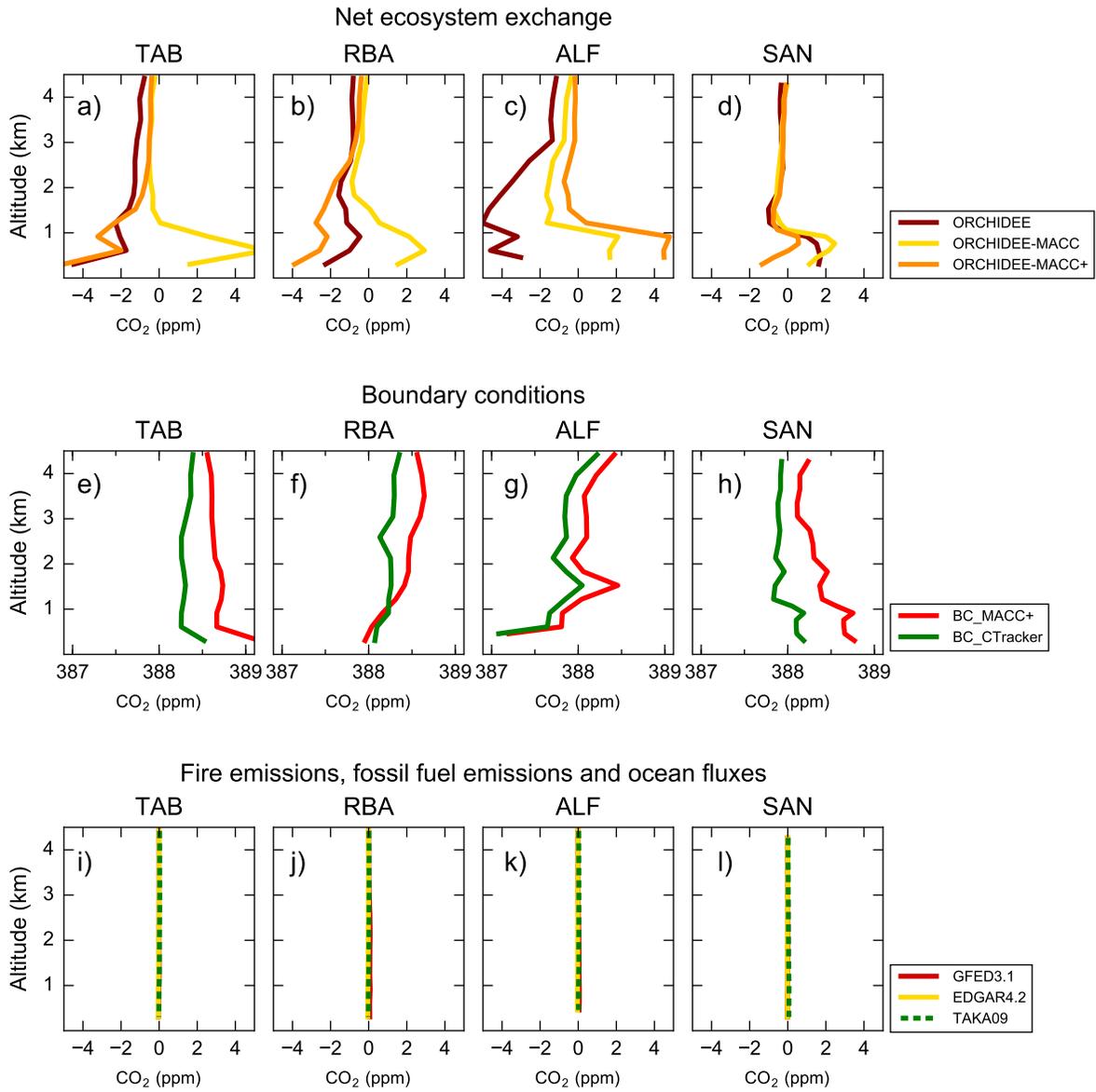


Figure 3.20: Sensitivity of seasonal mean vertical CO<sub>2</sub> profiles simulated with CH-BRAMS to (a – d) NEE, (e – h) boundary conditions and (e – h) fire and fossil fuel emissions and ocean fluxes for November – June (wet season) of 2010.

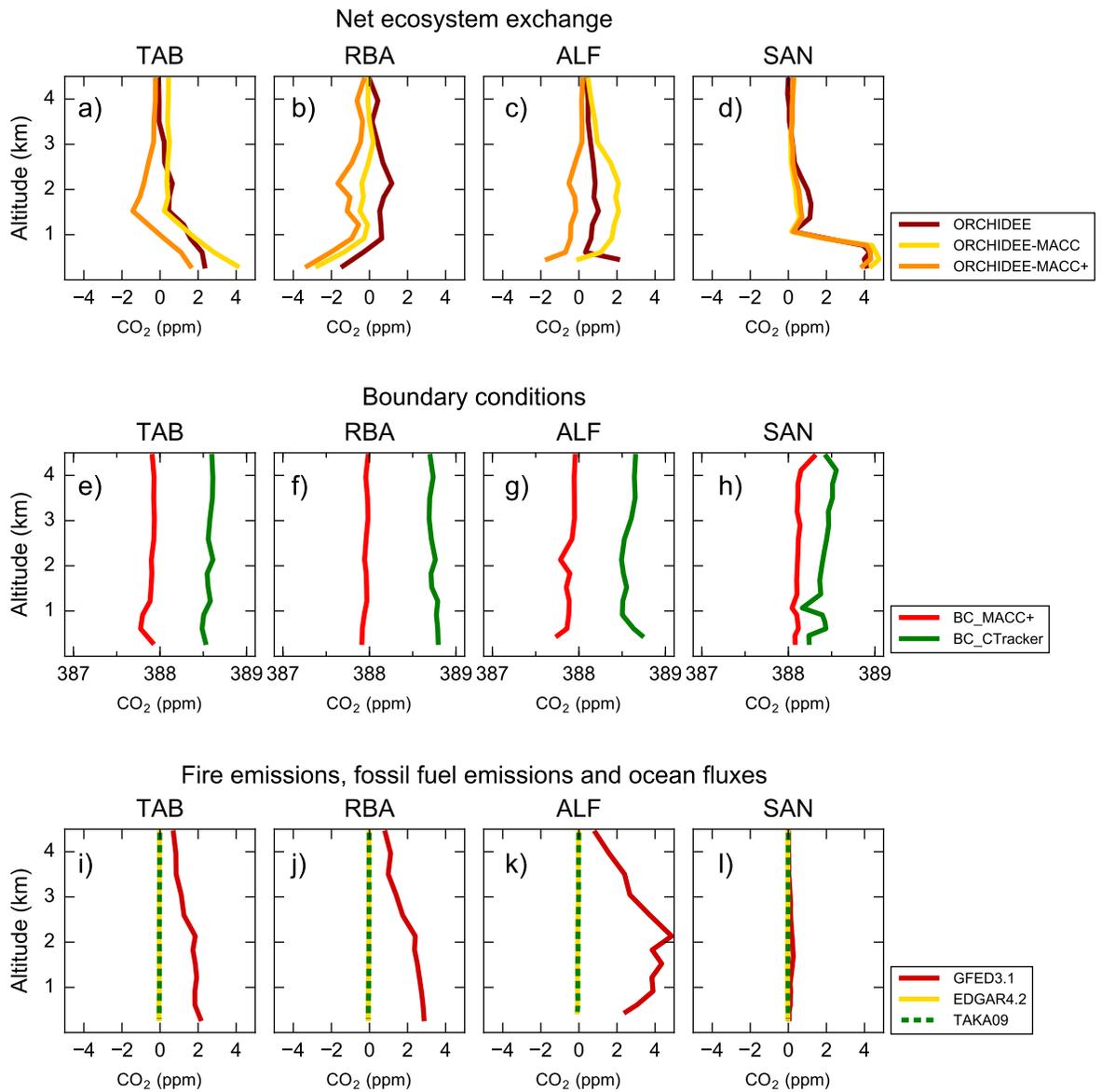


Figure 3.21: Sensitivity of seasonal mean vertical CO<sub>2</sub> profiles simulated with CH-ECMWF to (a – d) NEE, (e – h) boundary conditions and (e – h) fire and fossil fuel emissions and ocean fluxes for July – October (dry season) of 2010.

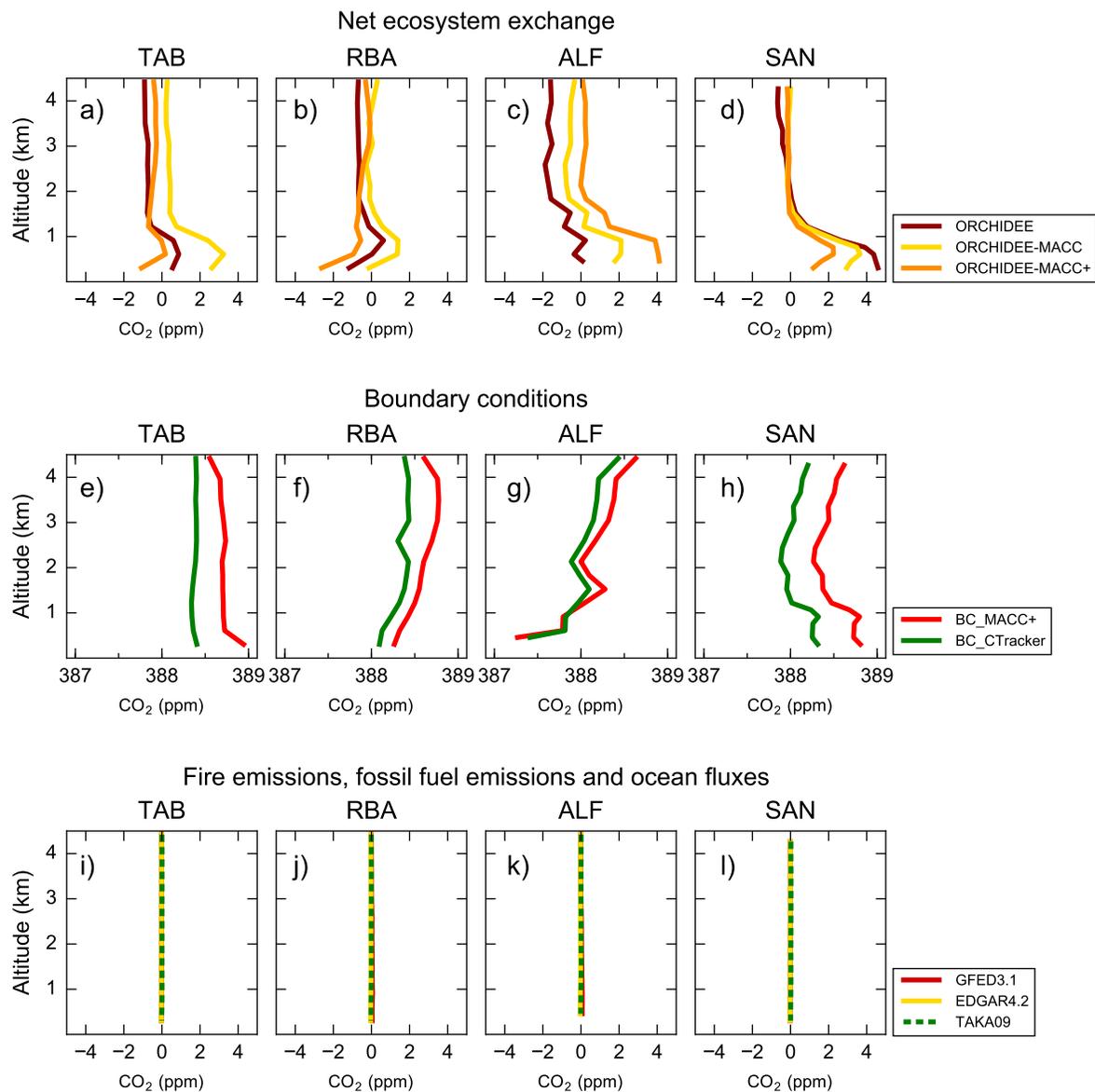


Figure 3.22: Sensitivity of seasonal mean vertical CO<sub>2</sub> profiles simulated with CH-ECMWF to (a – d) NEE, (e – h) boundary conditions and (e – h) fire and fossil fuel emissions and ocean fluxes for November – June (wet season) of 2010.

### Analysis of the gradient of CO<sub>2</sub> between the measurement sites

The horizontal gradients along the transects TAB–SAN, RBA–SAN and RBA–ALF (Figure 3.6) are shown in Figure 3.23. RMSE of the model-observation misfits varied up to 2 ppm across sites and seasons, and the correlation between observed and simulated gradients varied depending on the site, the season and the model. For instance, along TAB–SAN in the dry season, CH-BRAMS and CH-ECMWF exhibited correlations of 0.7 and  $-0.6$ , respectively, but along RBA–SAN in the wet season, both models showed correlations  $> 0.9$ .

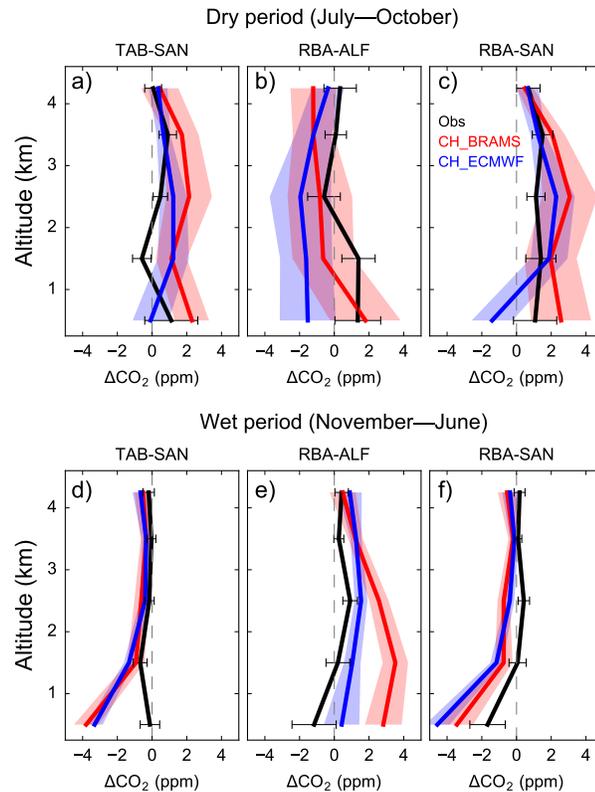


Figure 3.23: Observed and simulated mean seasonal CO<sub>2</sub> gradient between downwind and up-wind sites. Gradients are calculated for vertical bins every 1000 m from the surface to 5 km for (a – c) July – October (dry season) and (d – f) November – June (wet season) of 2010. Horizontal gradients were simulated with CH-BRAMS (red) and CH-ECMWF (blue). Horizontal bars: standard deviation of the observations. Shaded areas: standard deviation of the model simulations. CO<sub>2</sub> was simulated using ORCHIDEE NEE, EDGAR4.2 anthropogenic emissions, GFED3.1 biomass burning emissions, TAKA09 ocean fluxes, and BC\_MACC+ boundary conditions.

During the wet season, simulated gradients reproduced better the sign of the observed gradients within the first kilometer from the surface than during the dry season, except along TAB–SAN, where both models predicted a gradient of  $\sim -4$  ppm in the first layer, whereas the observations indicated an almost null gradient. Along RBA–ALF, CH-BRAMS showed much larger model-observation differences than CH-ECMWF. The former overestimated the gradient by as much as 4 ppm, up to four times than the latter. Likewise, CH-ECMWF underestimated the gradient RBA–ALF by  $\sim 3$  ppm during the dry season. During the dry season, results showed model-observation misfits of  $\sim 2$  ppm within the first 2 km,

larger than those during the wet season (except along RBA–ALF). Figure 3.23 shows that models' variability was slightly larger than that of the observations in the dry season, but it was rather well represented by the models in the wet season. Despite of the model-observation misfits, the observed gradients were small and models were able to capture their overall magnitude.

The difference between the two models ( $\Delta$ MODELS)—transporting the same set of fluxes—was, in general, higher within 0 – 2 km (up to 4 ppm) than above 2 km (typically  $\leq 1$  ppm).  $\Delta$ MODELS was higher during the dry season. Even with the same input fluxes, models predicted gradients of the same sign during the wet season, except along TAB–SAN (Figure 3.23d), but not in the dry season, which indicates important differences in the simulated transport.

Figures 3.24 to 3.27 show the gradients simulated with CH-BRAMS and CH-ECMWF, with different sets of NEE (panels a – c), and the contribution of the different fluxes and boundary conditions to the gradients (panels d – l) Along RBA–ALF during the wet season, for instance, the difference between ORCHIDEE-MACC+ and ORCHIDEE reached values  $> 5$  ppm within 0 – 2 km (Figures 3.25b and 3.27b). In addition, depending on the NEE, the sign of the gradients changed e.g. for CH-BRAMS along RBA–SAN during the dry season (Figure 3.24c), and for the three transects during the wet season (Figure 3.25(a – c)). Results also show that the assimilation of GA2014 in ORCHIDEE-MACC+ did not necessarily improve the fit to the observed gradients with respect to ORCHIDEE; misfits to observations were lower with ORCHIDEE-MACC, for instance, along the three transects during the wet season with CH-BRAMS (Figure 3.25(a – c)). Similarly, model-observation misfits with CH-ECMWF were lower with ORCHIDEE-MACC than with ORCHIDEE-MACC+ along TAB–SAN (Figure 3.26a) during the dry season and along all transects between pairs of aircraft sites in the wet season (Figure 3.25(a – c)). Note that for both models, the vertical structure of the simulated gradients and their average value were strongly influenced by the choice of NEE. The different NEE products tested here are very different and thus yielded very different concentrations gradients. The high uncertainty in NEE and the strong sensitivity of the gradients to these fluxes can thus explain the misfits between observed and simulated gradients (panels a – c in Figures 3.24 to 3.27), along with transport model errors.

The individual contribution of flux components shows that for either model and for both dry and wet seasons, there was always a small signature of ocean fluxes and fossil fuel emissions on the gradients, on the order of 1/100 of a ppm. The contribution of boundary conditions was one order of magnitude larger (typically  $\leq 0.7$  ppm). In contrast, I recall that NEE generated gradients of up to  $\sim 8$  ppm close to the surface. Note also the important influence of biomass burning emissions on the gradients during the dry season (panels a – c in Figures 3.24 and 3.26), was comparable to that of the NEE, with gradients of up to 3 – 4 ppm within 1 km, and up to  $\sim 2$  ppm above 2 km. The three transects exhibited an important impact from fire emissions, at least in the year 2010, as found by GA2014 based on gradients between the oceanic and the same profile sites. During the wet season, the contribution of biomass burning emissions was almost suppressed, limited to a few hundredths of a ppm. The signals of fossil fuel emissions and ocean fluxes on the gradients were nearly zero from 0 to 5 km, for both models during dry and wet periods.

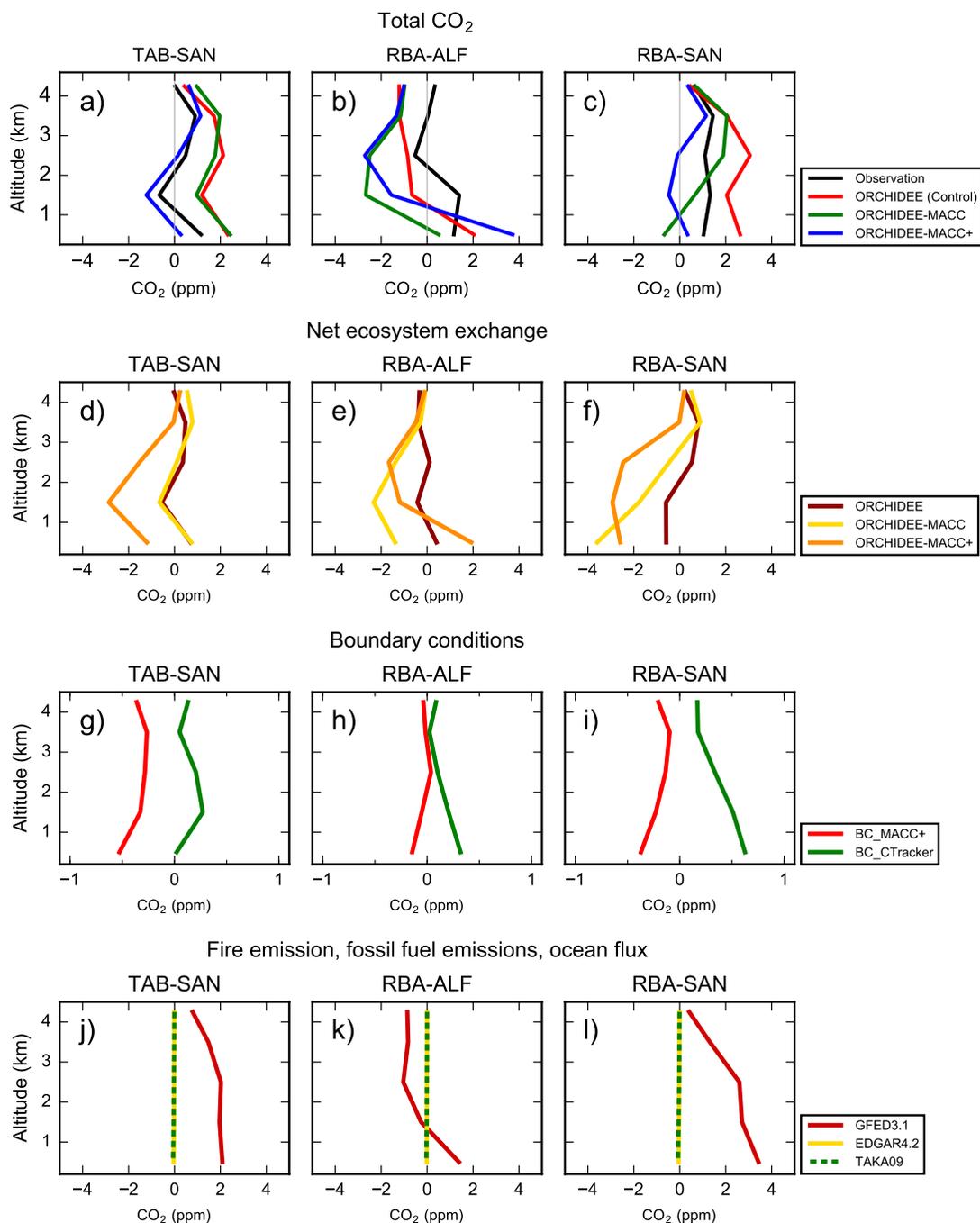


Figure 3.24: Observed and simulated mean seasonal CO<sub>2</sub> gradient between pairs of profile sites, and individual contribution from flux components and boundary conditions. Model simulations with CH-BRAMS for July – October (dry season). Gradients are calculated for vertical bins every 1000 m from the surface to 5 km.

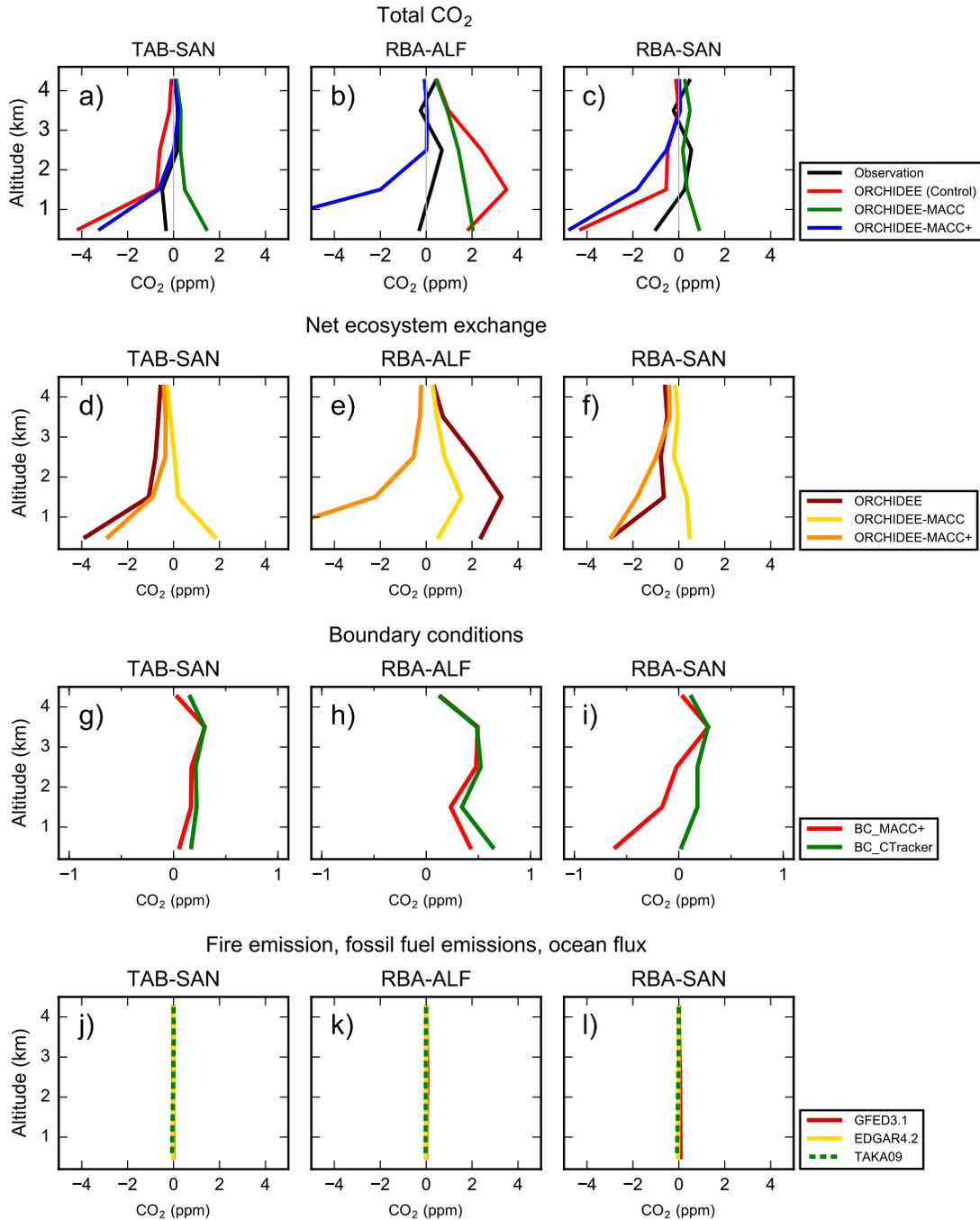


Figure 3.25: Observed and simulated mean seasonal CO<sub>2</sub> gradient between pairs of profile sites, and individual contribution from flux components and boundary conditions. Model simulations with CH-BRAMS for November – June (wet season). Gradients are calculated for vertical bins every 1000 m from the surface to 5 km.

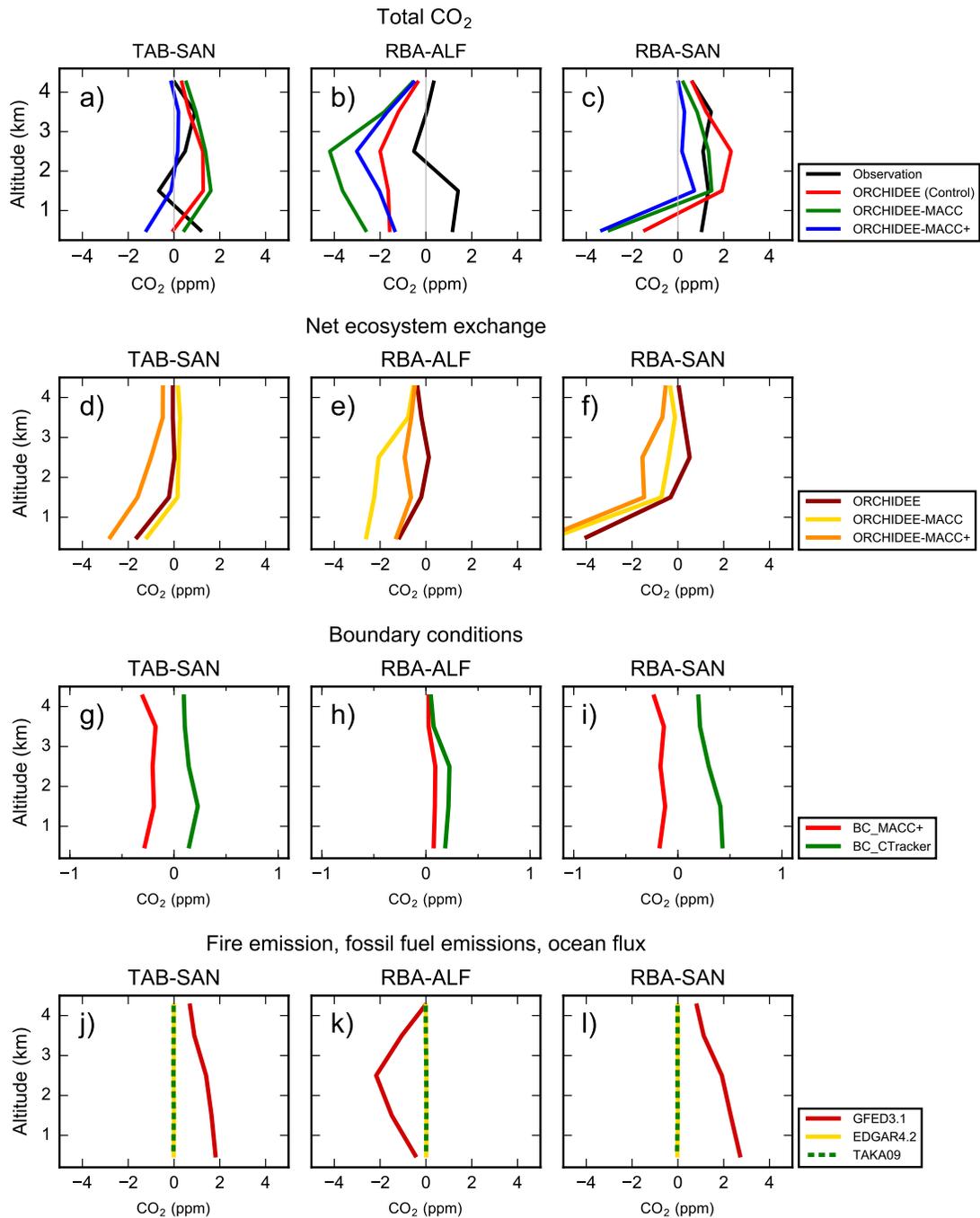


Figure 3.26: Observed and simulated mean seasonal  $\text{CO}_2$  gradient between pairs of profile sites, and individual contribution from flux components and boundary conditions. Model simulations with CH-ECMWF for July – October (dry season). Gradients are calculated for vertical bins every 1000 m from the surface to 5 km.

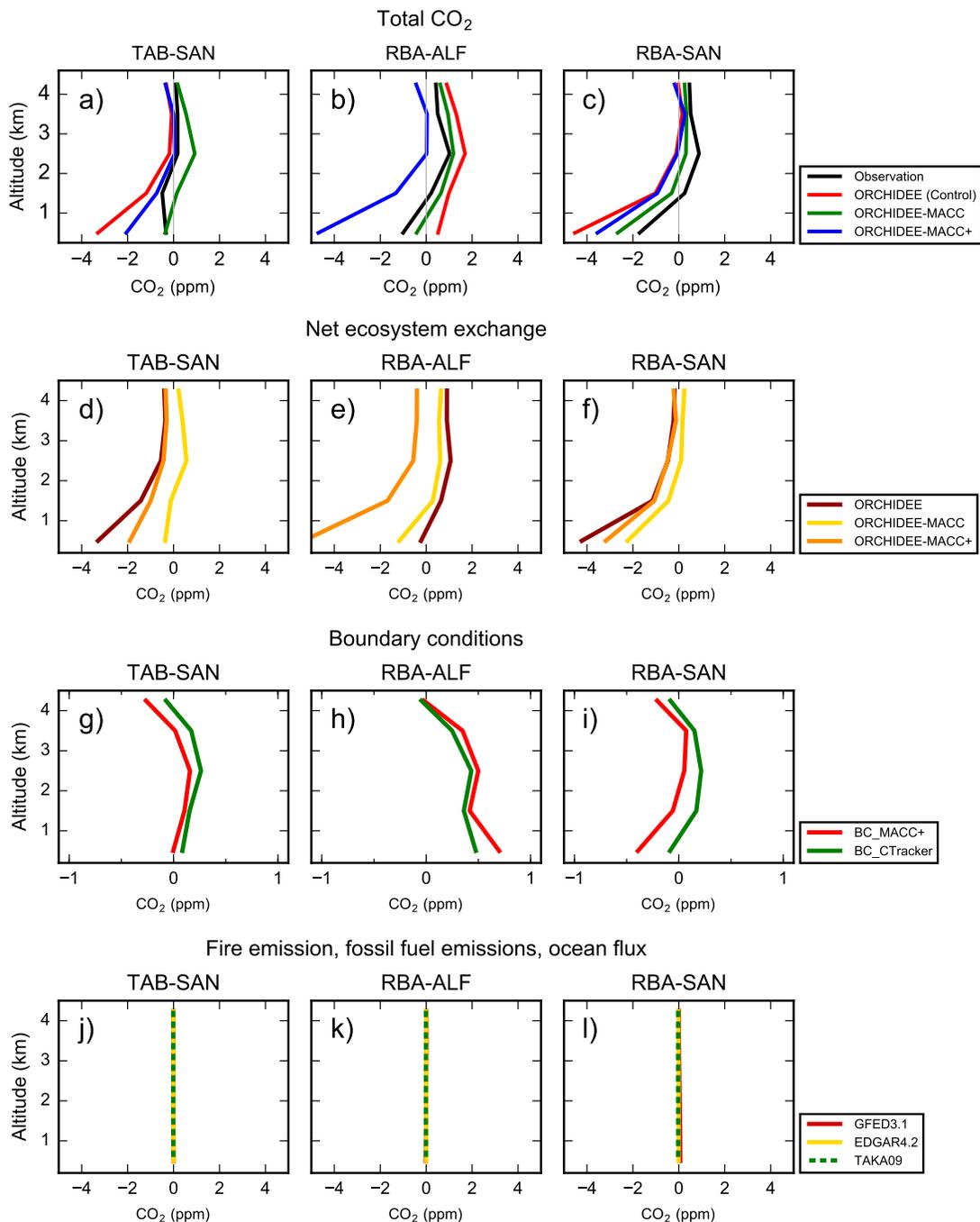


Figure 3.27: Observed and simulated mean seasonal CO<sub>2</sub> gradient between pairs of profile sites, and individual contribution from flux components and boundary conditions. Model simulations with CH-ECMWF for November – June (wet season). Gradients are calculated for vertical bins every 1000 m from the surface to 5 km.

Figures 3.28 and 3.29 show the observed and simulated vertical integral of the mean horizontal gradient along the three transects (dry and wet season, respectively), and the model simulations of the contribution of the flux components and boundary conditions to the vertical integral.

During the dry season, observed total CO<sub>2</sub> gradients, in absolute value, varied between ~0.4 and 1.0 ppm along the three transects, and reached about 0.2 ppm during the wet season. During both the dry and wet seasons, both transport models, in general, overestimated or underestimated the observed gradients, depending on the set of NEE used. Model-observation misfits reached up to ~3 ppm in the dry season, and up to ~2 ppm in the wet season. Further, only in one case (RBA-ALF, dry season, Figure 3.28b) the gradients simulated with the three NEE estimates agreed on the sign of the observed gradient.

Interestingly, during the dry season, models predicted positive and negative mean gradients due to biomass burning emissions along the transects (Figure 3.28(a – c)). Figure 3.30 depicts the location of the sampling sites in Amazonia relative to the mean biomass burning CO<sub>2</sub> emissions distribution from GFED3 during the dry period.

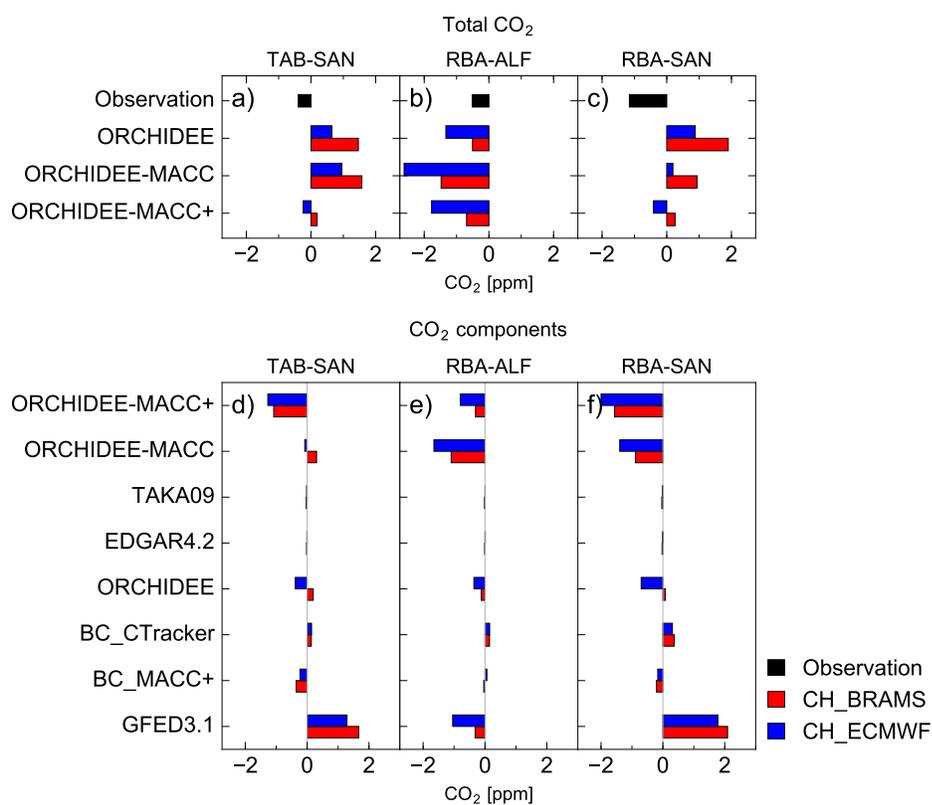


Figure 3.28: **(Top)** Observed and simulated integral of CO<sub>2</sub> horizontal gradients for 1000-m bins from the surface to 5 km asl, and **(bottom)** the individual contribution of NEE, biomass burning and anthropogenic emissions, ocean fluxes and boundary conditions to the simulated gradients. Gradients simulated with **(red)** CH-BRAMS and **(blue)** CH-ECMWF, for July – October (dry season).

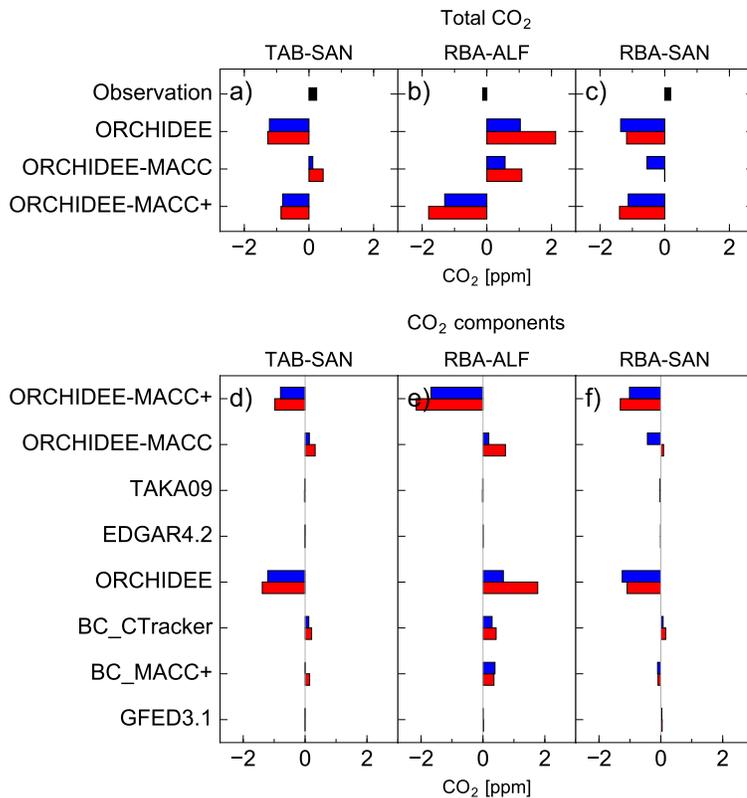


Figure 3.29: **(Top)** Observed and simulated integral of CO<sub>2</sub> horizontal gradients for 1000-m bins from the surface to 5 km asl, and **(bottom)** the individual contribution of NEE, biomass burning and anthropogenic emissions, ocean fluxes and boundary conditions to the simulated gradients. Gradients simulated with **(red)** CH-BRAMS and **(blue)** CH-ECMWF, for November – June (wet season).

In an independent experiment, using CH-BRAMS for illustrative purposes, biomass burning emissions were allowed to release CO<sub>2</sub> during 7 days, after which the emission stopped and the transport was allowed to continue for another 10 days. Figure 3.31 shows the signature of biomass burning emissions on the mean total CO<sub>2</sub> column in response to the 7-day emission event. Figure 3.31 shows that air circulation transported CO<sub>2</sub> northwest from fire emission zones around RBA and ALF (Figure 3.30). Along transects TAB–SAN and RBA–SAN CO<sub>2</sub> accumulated about the sites downwind of SAN, where fire emissions are close to zero, producing a positive downwind-upwind gradient (Figure 3.28). Figure 3.30 shows that ALF and RBA are close to emissions zones. Stronger emissions around ALF than around RBA might have created a negative gradient even though we could have expected that, on the first order, an accumulation of CO<sub>2</sub> from biomass burning emissions transported from ALF to RBA would have created a positive gradient along RBA–ALF.

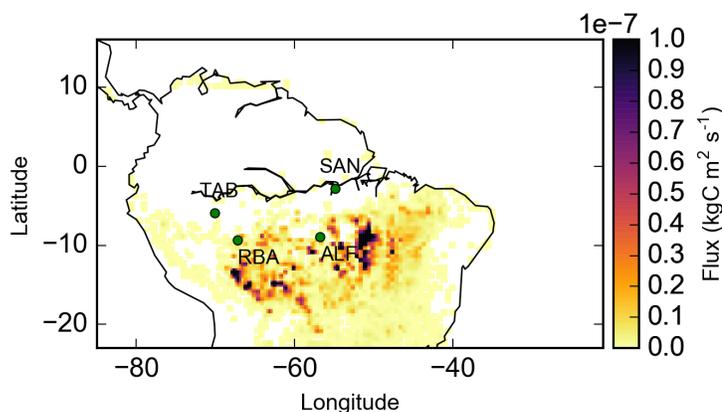


Figure 3.30: Mean biomass burning CO<sub>2</sub> emissions during July – October 2010 (dry period) from GFED3.1. **Green dots** indicate the location of local stations.

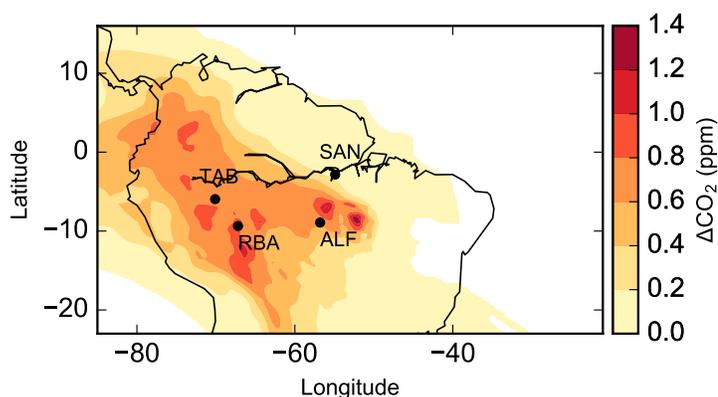


Figure 3.31: Mean total CO<sub>2</sub> column for August 11 – 29 2010, in response to an 7-day event of biomass burning emission during August 11 – 18, simulated with CH-BRAMS. Biomass burning emission were obtained from GFED3.1. **Black dots** indicate the location of local stations.

### 3.3 Discussion

In my comparison of the meteorology generated with the regional atmospheric model BRAMS and the global high resolution model ECMWF with radiosonde data, both models had similar statistical performance. The deviations from observed values are comparable to those found in previous studies. In their evaluation of a former version of BRAMS, based also on radiosonde data, [Freitas et al. \(2009\)](#) reported  $STD_{MISFIT}$  values for potential temperature at the surface of  $\sim 2\text{K}$ , and  $\sim 1\text{K}$  at higher altitudes. Results in [Figures 3.8 and 3.7](#) give  $STD_{MISFIT}$  values of  $1.1 - 1.8\text{K}$  at the surface and  $0.7 - 1.1\text{K}$  for higher levels. For wind speed, [Moreira et al. \(2013\)](#) obtained mean RMSE values of  $\sim 1.5 - 3.5\text{m s}^{-1}$  using the information of 11 radiosonde stations distributed across tropical South America, and their results are comparable to the results obtained in [Section 3.2.1](#). [Bela et al. \(2015\)](#) carried out simulations with the models BRAMS and WRF-Chem and evaluated them at Manaus against radiosonde data, and found mean model-observation misfits of up to  $\sim 10\text{K}$  for dew point temperature and  $> 100^\circ$  for wind direction. My simulations at

Manaus indicated mean differences of approximately 4 – 100° and 5 – 91° for wind speed, for ECMWF and BRAMS, respectively, comparable to the study of [Bela et al. \(2015\)](#). For dew point temperature, differences to the observations also reached up to ~10 K for both models, comparable to [Bela et al. \(2015\)](#), as well.

Comparison to weather station data suggests that both models reproduced surface atmospheric conditions as well as in previous studies focusing on atmospheric dynamics in the Amazon. RMSE values of 1.7 – 2 K for potential temperature, 1.5 – 1.9 K for dew point temperature and ~1.5 m s<sup>-1</sup> for wind speed are comparable to the results of [Bela et al. \(2015\)](#), who obtained RMSEs of 1.7 – 2.8 K, 1.7 – 2.7 K and 1 – 1.4 m s<sup>-1</sup> for the same variables, respectively. Although RMSE values for BRAMS were 10 – 18% higher than for ECMWF, values over the two transition periods had the same order of magnitude for all the aforementioned variables. RMSE for precipitation for both models was similar (~1.9 mm), and the bias indicates that both models overestimated precipitation by < 1 mm. Larger RMSE and bias were found by [Moreira et al. \(2013\)](#). Note that ECMWF is an operational forecast system that assimilates data from multiple data sets, whereas my simulations with BRAMS did not use an assimilation package.

Comparison to TRMM precipitation data suggests that the large-scale distribution of precipitation across tropical South America during austral summer and winter (July – September and January – March, respectively) was well represented by both models. Both models captured the large-scale patterns of the ITCZ and SACZ. Yet mean precipitation rates during sampled periods (January and July 2010) were underestimated by both models, which may have repercussions on the intensity and distribution of the convective transport—critical for CO<sub>2</sub> transport. Overall, confrontation of simulated meteorology with in-situ and remote sensing observations indicates reasonable performance of both models for meteorology; it would be expected that both atmospheric models simulate the wind fields and convection in a realistic manner.

My comparison of observed and simulated mean CO<sub>2</sub> vertical profiles, however, revealed important differences with the observations from both transport models. These differences cannot be attributed to transport errors only, since the prescribed fluxes can also be erroneous. However, the strong differences between the CO<sub>2</sub> simulated with CH-BRAMS and CH-ECMWF ( $\Delta$ MODELS) revealed a strong impact from transport modeling uncertainties—fundamentally meteorological modeling errors—on the simulation of the CO<sub>2</sub> vertical profiles. Model-observation differences in CO<sub>2</sub> of ~0.1 – 6.7 ppm were identified within the PBL. Observed BL–FT differences at the sampling sites (Table 3.5) were > 0 during the dry season. CH-BRAMS, however, underestimated the magnitude of the gradients, and even simulated negative gradients, opposite to the observations, even though fire emissions were included in the surface fluxes. Conversely, CH-ECMWF simulated the right sign of the gradients, but overestimated the magnitude (except at RBA). During the wet season, observed BL–FT gradients were < 0 at all sampling sites. It was CH-ECMWF that simulated positive gradient at all sites but one (RBA). CH-BRAMS, on the other hand, failed to reproduce the right sign of the gradient only at SAN, but underestimated the magnitude at sites. Surprisingly, transport simulations with both CH-BRAMS and CH-ECMWF did not show a significant sensitivity to the deep convection parameterization that explained the model-observation misfits. Using a former version of CH-BRAMS, [Herrmann and Freitas \(2011\)](#) evaluated the model sensitivity to deep convection. Their results showed the strongest effect of deep convection at altitudes as high as 14 km, where maximum detrainment by deep convection occurs. Unfortunately, no observations were available at that altitude to test the effect of this transport mechanism.

The analysis of the contribution of the different types of flux in the region illustrated the impact of the uncertainties in these different fluxes on the whole CO<sub>2</sub> vertical profiles. The signature of NEE on CO<sub>2</sub> was dominant in the four observation sites evaluated. In general, the signal was strongest close to the surface (i.e. 0 – 2 km). The second strongest flux influencing CO<sub>2</sub> was biomass burning emissions during the dry season, which had a signal comparable to that of NEE (depending on the NEE used) both close to the surface and in altitude. The third important component influencing CO<sub>2</sub> was the boundary conditions. The sensitivity of both individual measurements in profiles and horizontal gradients to boundary conditions is smaller than those of NEE or fire emissions at all altitudes. Further, in the gradients the signal of boundary conditions was close to zero. As to ocean fluxes and fossil fuel emissions, their impact on both vertical profiles and gradients virtually zeros.

All the indicators related to model-observation CO<sub>2</sub> misfits (vertical profiles and horizontal gradients between pairs of aircraft sites) showed a dominant sensitivity to the choice of NEE. In some instances, simulated profiles or horizontal gradients were closer to the observations with ORCHIDEE-MACC fluxes than with ORCHIDEE-MACC+, even though ORCHIDEE-MACC+ assimilated the GA2014 profiles in a global inversion. This indicates an inconsistency between NEE from this global inversion and the CO<sub>2</sub> vertical profiles when using high-resolution transport. In a sense, this result proves that regional transport models produce a very different CO<sub>2</sub> concentration field than a global model, so that when an inversion will be applied with regional models, it will produce a different NEE when assimilating the GA2014 profiles.

The ratio of the uncertainty in NEE to that in other fluxes, boundary conditions or transport error, as a function of the observation type (individual measurements or horizontal gradients), provided an indicator of the potential influence of uncertainties (in NEE or other components) in CO<sub>2</sub> inversions. Differences between NEE estimates ( $\Delta$ NEE) provided a measure of uncertainty on NEE. Likewise, differences between BC\_MACC+ and BC\_CTracker ( $\Delta$ BC) served as a proxy of the uncertainty in boundary conditions. Horizontal gradients at different altitudes, close to the surface or the vertical integral showed higher  $\Delta$ NEE: $\Delta$ BC ratios than individual measurements in vertical profiles.

For biomass burning emissions, I did not evaluate a second product to assess the typical uncertainty in these emissions. I assumed that a strong signature of fire emissions was likely associated to large uncertainty and conversely. Thus, I used the ratio of the signature of NEE to that of fire emissions as a proxy of the  $\Delta$ NEE: $\Delta$ FIRE ratio. When analyzing individual measurements or gradients between pairs of measurement sites,  $\Delta$ NEE: $\Delta$ FIRE ratios were similar, close to the surface or in altitude. Therefore, it is unclear that gradient-based observations reduce the uncertainty in fire emissions with respect to the uncertainty in NEE more than individual observations in profiles. Among the gradient-based observations, however, the horizontal gradients close to the surface showed the highest  $\Delta$ NEE: $\Delta$ FIRE values.

Misfits between the models ( $\Delta$ MODELS) were lower in gradients between pairs of sites than in individual measurements, but depending on the transect and the NEE product used,  $\Delta$ MODELS can be as large as  $\Delta$ NEE either close to the surface or in altitude. Thus, transport error is still important even in the horizontal gradients. Further,  $\Delta$ NEE: $\Delta$ MODELS ratios were similar in both individual measurements and gradient-based observations, so that it is not clear whether profiles or gradients reduce the most the transport error vs. the uncertainty in NEE. Horizontal gradients close to the surface showed model misfits larger than the difference between NEE estimates. Transport error and NEE differences in the gradients decreased with altitude, but  $\Delta$ NEE: $\Delta$ MODELS ratios were higher in altitude

than close to the surface, so that in altitude transport error does not mask the uncertainty in NEE as it does close to the surface.

The high sensitivity of the different observation types to NEE relative to other flux components or boundary conditions, i.e. the fact that observed gradients are sensitive dominantly to NEE first, then to biomass burning emissions during the dry season and transport error in second, and lastly to boundary conditions and other fluxes, is a positive sign regarding the potential of the NEE inversion. However, separating the signals of NEE from that of biomass burning emissions during the dry season, and from that of transport modeling errors in an inversion system seems challenging.

## 3.4 Conclusions

Accurate representation of atmospheric transport is critical to CO<sub>2</sub> inversion studies. Transport modeling errors associated with the coarse resolution of global transport models influence (to a poorly known extent) the Amazon CO<sub>2</sub> fluxes inferred using global inversions. Regional atmospheric models can potentially improve inversion results over the region thanks to a higher resolution, allowing the models to capture some regional and mesoscale processes, and to use higher resolution topography, coastline and flux input maps. Airborne measurements of atmospheric CO<sub>2</sub> are a useful tool for evaluation of CO<sub>2</sub> transport and should, in principle, provide valuable information about the carbon mass balance in Amazonia, and thus provide a strong constraint in NEE inversions.

However, the modeling of quasi instantaneous CO<sub>2</sub> vertical profiles approximately every two-weeks within the Amazon basin, where vertical transport is complex, and with a high influence from biomass burning CO<sub>2</sub> emissions appears to be challenging as shown by the results from this chapter. Here, I aimed at evaluating the benefit of the regional meteorological model BRAMS—which has the advantage of being calibrated from previous studies for atmospheric studies in Amazonia, a higher resolution, and being coupled to a more advanced land surface scheme—relative to the global system ECMWF—which has the advantage of assimilating weather satellite, radiosonde and surface stations data at relatively high resolution—when both models provide meteorological fields for tracer transport with the offline tracer transport model CHIMERE. Both BRAMS and ECMWF simulated the meteorology in tropical South America with similar performance in terms of the statistics of the model-observation misfits, but the simulations were very different between them. Yet results from both models were comparable to previous atmospheric model evaluations in this region.

CO<sub>2</sub> transport simulations with both model configurations also showed similar performance, with large misfits to observations, both for individual measurements and when considering overall shape of the observed vertical profiles. Rather than demonstrating the superiority of the purely regional, transport-meteorological configuration CH-BRAMS, the comparisons between CH-BRAMS and CH-ECMWF highlights and characterizes the high impact of transport modeling uncertainty, even when using state-of-the-art model configurations for modeling such quasi instantaneous vertical profiles of CO<sub>2</sub>. Therefore, both meteorological forcings will be used in Chapter 4 for the regional inversions, empirically propagating the transport model errors on the estimates of NEE from these inversions.

In a general way, the tests of sensitivity to meteorological forcing or fluxes did not really reveal a “best configuration” for simulating such profiles. The comparison between the sensitivity to the meteorological forcing and the sensitivity to the different types of CO<sub>2</sub>

fluxes was rather used to provide insight about the ability for a regional inverse modeling system to filter the uncertainty in the NEE from the other error sources when simulating CO<sub>2</sub>. I made sensitivity analyses of the CO<sub>2</sub> signal across four types of observation vectors: the individual measurements, the horizontal gradients between pairs of measurement profiles for five vertical layers, the horizontal gradients in the surface layer between pairs of profiles sites, or horizontal gradients between the vertically integrated profiles. In all cases, the observations were mostly sensitive to NEE, and then to fire emissions during the dry season and to the transport models. The sensitivity to boundary conditions was much smaller, while the sensitivity to fossil fuel emissions and ocean fluxes was nearly negligible. This general result suggests that there could be a high potential for adjusting the NEE based on the CO<sub>2</sub> data from the aircraft profiles without much impact from uncertainties in the other types of fluxes or in the transport. However, the analysis also revealed that it is very difficult to attribute specific patterns of the model-observation misfits to transport modeling errors, or to NEE or biomass burning emissions during the dry season, which may limit the skill of the regional atmospheric inversion to invert NEE based on the assimilation of the CO<sub>2</sub> vertical profiles.

At least, the analysis of different types of observations that could be assimilated in the inversion (individual measurements or gradients between the vertical profiles) demonstrates that in all cases the inversion should not be influenced by uncertainties in the fossil fuel emissions or in the ocean fluxes, and that the use of horizontal gradients should nearly cancel the impact from boundary conditions, which is significant when analyzing individual concentration measurements. Gradients of CO<sub>2</sub> between pairs of sites also showed lower misfits between CH-BRAMS and CH-ECMWF than the individual CO<sub>2</sub> profiles, either close to the surface or in altitude.

In horizontal gradients close to the surface, differences between CH-BRAMS and CH-ECMWF were larger than the differences between NEE estimates. Transport error and NEE differences in the gradients decreased with altitude, and transport error became less important with respect to the uncertainty in NEE in altitude than close to the surface, so that in altitude the transport error should not mask the uncertainty in NEE as much as it does close to the surface. Therefore, there could be a gain from assimilating the gradients at higher altitudes in comparison to assimilating gradients close to the surface only, or to assimilating vertically integrated horizontal gradients. As to fire emissions, the uncertainty in these emissions seems to have a similar weight with respect to that in NEE across the four observation types. Yet horizontal gradients close to the surface showed the lowest influence of biomass burning emissions with respect to NEE.

Therefore, in a regional inversion for South America, the assimilation of CO<sub>2</sub> gradients, especially considering gradients for different vertical layers from the surface to high altitude, seems to offer some stronger potential, relative to the assimilation of individual measurements in vertical CO<sub>2</sub> profiles, to filter the signal of NEE from that of boundary conditions, and to a lesser extent, from transport modeling errors. However, regarding the influence of biomass burning emissions, and for a large part, of transport modeling errors, there is no clear indication that assimilating horizontal gradients instead of individual measurements in vertical profiles should strengthen the potential of the inversions. Given that defining gradients between pairs of profile sites strongly decreases the number of data that can be assimilated, the approach may not be more robust than using individual data. Inversion experiments in the next chapter, using each observation type separately, should provide further information on the consistency and robustness of the results derived when using these different types of observation vectors.



---

## Regional atmospheric inversion of CO<sub>2</sub> in Amazonia

This chapter is devoted to the development and application of a dedicated regional atmospheric inversion modeling framework to estimate the net ecosystem exchange (NEE) in the Amazonia. This inversion framework uses the two transport models presented in Chapter 3 to optimize the NEE in the Amazon using CO<sub>2</sub> measurements from aircraft profiles collected by [Gatti et al. \(2014\)](#) (GA2014, henceforth).

The previous chapter showed the need for analyzing the sensitivity of such an inversion to the atmospheric transport model configurations developed and/or evaluated in this thesis, or, to be more precise, to the meteorological product used to force the regional tracer transport model CHIMERE. The analyses of Chapter 3 also indicated that observations consisting of horizontal CO<sub>2</sub> gradients between pairs of aircraft sampling sites along the main wind direction, instead of individual CO<sub>2</sub> measurements at each vertical profile, increased the ability to distinguish the atmospheric signature of the uncertainty in the NEE by reducing that of the uncertainty in the CO<sub>2</sub> boundary conditions. The improved ability to isolate the atmospheric signature of the uncertainty in the NEE should strengthen the precision of the NEE inversion. However, such an improvement may be limited by the fact that the sensitivity to transport errors and to the signature of biomass burning CO<sub>2</sub> ( $E_{\text{FIRE}}$ ) emissions in the dry season did not decrease in the horizontal gradients along the dominant wind direction compared to individual CO<sub>2</sub> measurements. A cautious selection of wind conditions for the selection of CO<sub>2</sub> gradients between the aircraft profiles imposes a stringent reduction of the number of data to be assimilated ([Bréon et al., 2015](#); [Staufner et al., 2016](#)). These conclusions from the previous chapter support to carry out inversions with different strategies for the definition of the observation vector. Therefore, in this chapter I have tested the assimilation of individual CO<sub>2</sub> measurements in vertical profiles at the different measurement sites and horizontal CO<sub>2</sub> gradients along the wind direction between these profile sites—close to the ground, for different vertical bins or considering the total CO<sub>2</sub> columns—using the CH-BRAMS and CH-ECMWF transport models. The prior NEE was obtained from a vegetation model simulation and it has been improved to include an estimate of land use change (LUC) CO<sub>2</sub> emissions. As shown in Chapters 2 and 3, the statistical structure of the uncertainty in such a prior estimate in Amazonia is particularly difficult to assess.

In such an exploratory context, and given all uncertainties regarding the right config-

uration of the regional inversions, I developed an analytical inverse modeling framework with few control parameters rather than a system which could solve for the fluxes at relatively high resolution (Chapter 2). Analytical inverse modeling frameworks with a small number of control variables allow enable us to carry out a large number of tests with different inversion parameters or observational datasets at low computational cost. Furthermore, a large number of statistical diagnostics, which are very useful for understanding the inversion problem, limitations, opportunities and possible adaptations can be easily derived from such analytical systems.

Thus, I developed a regional analytical inversion system for tropical South America (Section 4.1). The set of observations from GA2014 now covers the period 2010 – 2012 (see Alden et al., 2016), but in order to evaluate the system without multiplying the computations, and given the results and conclusions I obtained in that period (Chapter 3), I only tested the inversion for two trimesters in 2010 during the dry and the wet seasons. As in the previous chapter, these two contrasting periods, with the enhanced and prolonged dry conditions during the dry season in 2010 and the more humid conditions of the wet season—despite the wet season of 2010 also affected by el El Niño (Marengo et al., 2011)—should allow for the detection of the impact of extreme climate conditions on NEE. In principle, the impact of such conditions on the NEE are hardly caught by the prior NEE estimates available for this thesis (see Chapters 2 and 3) and therefore, these time periods offer an interested test of the ability of the inversion to detect and quantify this impact, and more generally to improve the seasonal and inter-annual variability of the NEE in the Amazon basin.

## 4.1 Inverse modeling setup

I built an analytical inversion system with the two transport models presented in Chapter 3. Using two different transport models allows for a coarse estimate of the effect of transport uncertainties on optimized fluxes. The inversions covered the periods of January to March (JFM) in the wet season, and July to September (JAS) in the extreme dry season of 2010. The simulation domain is the one defined in Chapter 3 and that extends between 16.3° N – 23.1° S and 85.1° W – 19.7° W (Figure 4.1).

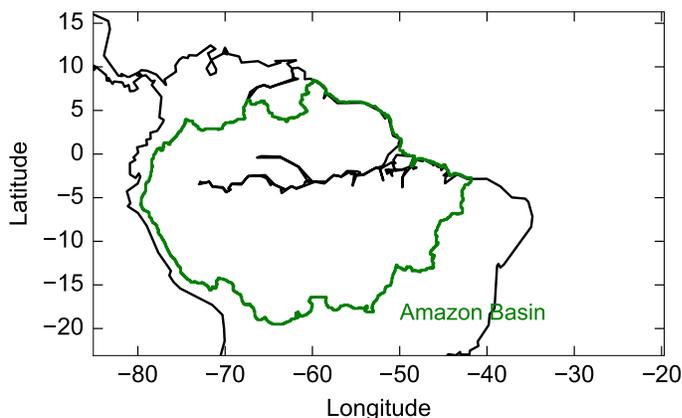


Figure 4.1: Geographic domain for the inversion experiments.

### 4.1.1 Control vector

In the analytical approach chosen here, building the observation operator matrix requires to perform a transport simulation for each control parameter, which is computationally demanding when a large number of control parameters are sought, e.g. solving for fluxes on a large number of grid-cells (Section 1.1.3). Therefore, fluxes were spatially and temporally aggregated so that each control parameter corresponded to the budget of a type of flux for a major geographic subdivision of the modeling domain. The time aggregation is one week to one month (depending on the type of flux), i.e. weekly to monthly mean budgets of the fluxes are optimized, distinguishing between day- and nighttime fluxes in the case of NEE. NEE and  $E_{\text{FIRE}}$  emissions were optimized independently based on different geographical subdivisions of tropical South America. While this discretization reduces the number of control parameters, it may introduce aggregation errors (i.e. misfits between the model and the  $\text{CO}_2$  measurements due to uncertainties in the flux patterns within a control region and temporal window) that influence the precision of the estimated fluxes (Kaminski et al., 2001). For a given flux type, the subdivision of South America for the definition of the corresponding control parameters was chosen so that each type of flux was relatively homogeneous within a given region, or, at least, so that the variations of this type of flux within the region did not influence the concentrations at the measurement locations. This should limit the impact of uncertainties in the distribution of the fluxes within the control regions—which is not modified by the inversion—on the concentrations at the measurement locations, and thus limit the aggregation errors and their projection into the inverted fluxes.

I solved for weekly or monthly flux budgets for a set of 2 to 8 regions (see the next section for how they were chosen following the general strategy defined above), depending on the flux type, for two January – March (JFM) and July – September (JAS) in 2010. I estimated separately (a) the NEE (separating day- and nighttime fluxes) defined here as the net  $\text{CO}_2$  exchange between the terrestrial surface and the atmosphere, excluding  $E_{\text{FIRE}}$ , (b)  $E_{\text{FIRE}}$  emissions, and (c) the monthly budget of ocean  $\text{CO}_2$  fluxes.

### 4.1.2 Flux products used in the inverse modeling framework and definition of the control regions

Different sources of information were used to provide prior estimates of the flux budgets for the different weeks/months and regions corresponding to the control vector (Table 4.1). These datasets provided the flux distribution within the control sub-regions and time windows (as part of the observation operator; see Section 4.1.4), and the fluxes that were not controlled by the inversion (Table 4.1) (also a component of the observation operator). Both controlled and prescribed fluxes are discussed below. The spatial distribution of the mean seasonal surface fluxes from these products for NEE (including LUC emissions; see Land use change  $\text{CO}_2$  emissions below),  $E_{\text{FIRE}}$  emissions and ocean fluxes is illustrated in Figure 4.2. Most of them corresponded, or very similar, to the flux estimates used in Chapter 3. Note that neither the spatial nor the temporal distributions given by these fluxes maps within the control sub-regions or time windows were controlled by the inversion.

#### Annually balanced net ecosystem exchange

I used an estimate of the NEE calculated as the difference between ecosystem respiration (ER) and gross primary productivity (GPP) provided by a simulation of the ORCHIDEE

Table 4.1: Datasets used to derive the prior flux budgets and prescribe the spatial and temporal flux distribution (not controlled by the inversion) within the control sub-regions/weeks (as part of the observation operator).

Flux	Data source	Number of regions	Optimization
Net ecosystem exchange	Difference between the ecosystem respiration and gross primary productivity (ER–GPP) simulated with the ORCHIDEE model at 2° resolution every 3 hours. Corrected to account for land use change CO <sub>2</sub> emissions.	8	Yes (regional, weekly)
Biomass burning emissions	Monthly total CO <sub>2</sub> emissions from the Global Fire Emissions Database version 4 (GFED4, <a href="#">Giglio et al., 2013</a> ) at 0.25° horizontal resolution	2	Yes (regional, weekly)
Air-sea CO <sub>2</sub> exchange	Annual mean CO <sub>2</sub> flux from the climatology of <a href="#">Takahashi et al. (2009)</a> .	2	Yes (regional, monthly)
Fossil fuel emissions	Annual emissions from EDGAR 4.2 FT database at 0.1° resolution.	1	No

model ([Krinner et al., 2005](#)). ORCHIDEE simulated ER and GPP at 2° spatial resolution and with a 3-hour time step, driven with meteorological fields from CRUNCEP reanalysis, following the TRENDY protocol (see [Sitch et al., 2015](#)). This source NEE field—the same as used in Chapter 3—was interpolated to the transport model grid (~35 km). ORCHIDEE uses a single plant functional type for evergreen tropical forests, with the same model parameters, which ignores the diversity of forest phenology, structure, biomass and productivity of the Amazon biome, and thus provides spatially smoothed NEE fields, depending only upon regional climate differences. In this simulation, the NEE was annually balanced and did not account for disturbance due to land use change or wildfires. To account for LUC emissions, a correction was applied (see *CO<sub>2</sub> Emissions from Deforestation*) to this NEE dataset to provide the prior estimate of the NEE sub-regional budgets and of the NEE spatio-temporal distribution in the observation operator for the inversions.

For the NEE control areas I assumed that above-ground woody biomass productivity (NPP) was a proxy of the spatial heterogeneity of NEE. In Amazonia, NPP seems to be strongly influenced by soil fertility ([Aragão et al., 2009](#); [Malhi et al., 2004](#); [Quesada et al., 2012](#)) and correlated to forest biomass turnover. [Fiktau \(1969\)](#) proposed four fertility regions were proposed for Amazonia, based on geological and geomorphological soil characteristics. I used these regions, adapted from [Feldpausch et al. \(2012\)](#), as a basis for the definition of the NEE control areas. Each region was further divided into two sub-regions, for a total of eight NEE control sub-regions (Figure 4.3a), to decrease the influence of aggregation errors.

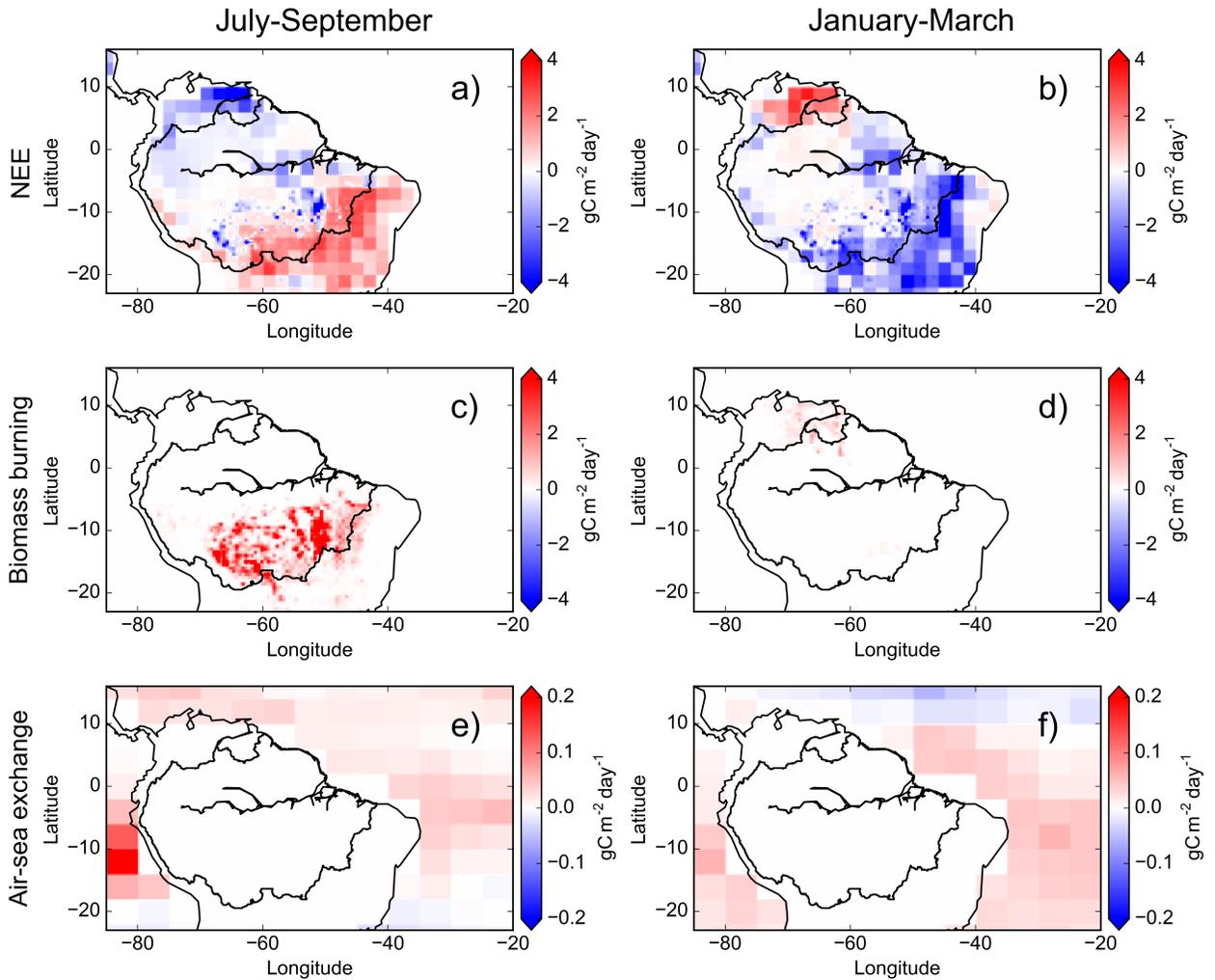


Figure 4.2: Maps of the mean seasonal (a, b) NEE (including LUC CO<sub>2</sub> emissions), (c, d) biomass burning CO<sub>2</sub> emissions and (e, f) air-sea CO<sub>2</sub> exchange used to derive the prior flux budgets and their spatio-temporal distribution within the observation operator for July – September (dry period) and January – March (wet period)

### Land use change CO<sub>2</sub> emissions

LUC CO<sub>2</sub> ( $E_{LUC}$ ) emissions, mainly due to deforestation, are an important component of the carbon balance of tropical South America. For instance, during the period 2005 – 2009 carbon emissions due to deforestation have been estimated at  $\sim 0.31 - 0.48$  PgC (Gloor et al., 2012). Therefore, I accounted for an estimate of  $E_{LUC}$  calculated as

$$E_{LUC} = NEE_{without\ LUC}^* + NEE_{without\ LUC}^* \quad (4.1)$$

Where  $NEE_{without\ LUC}^*$  and  $NEE_{without\ LUC}^*$  are monthly estimates of NEE with and without LUC, respectively. Both estimates were obtained from two simulations with the ORCHIDEE model at  $0.5^\circ$ , generated following the TRENDY protocol (see Sitch et al., 2015). Accord-

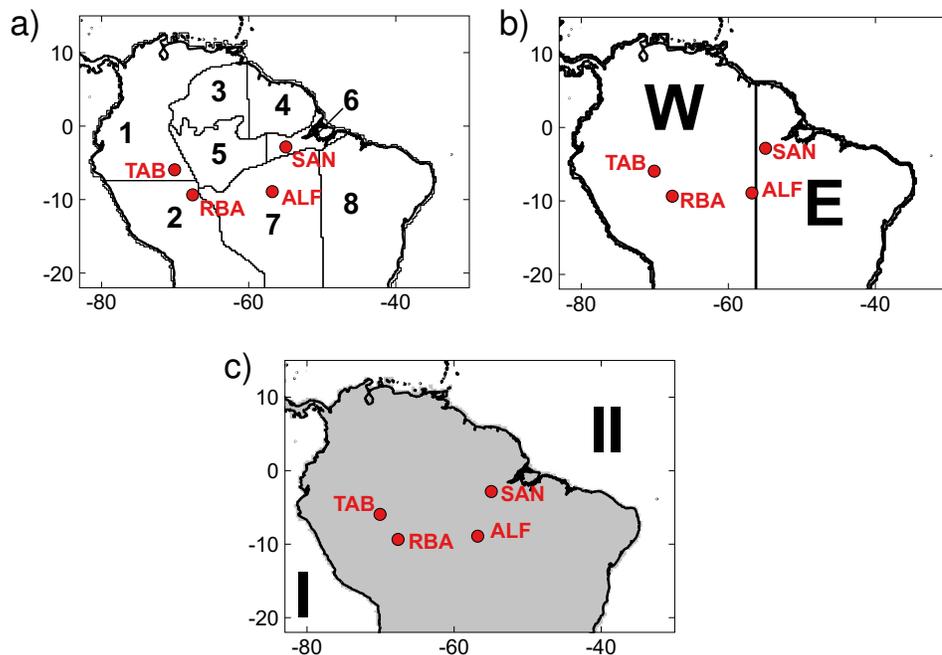


Figure 4.3: Regions for the optimization of (a) net ecosystem exchange (regions 1–8), (b) biomass burning emission (regions E and W) and (c) ocean fluxes (regions I and II).

ing to this protocol, model simulations covered the period 1861 – 2010. To generate the model is forced by annually variable vegetation maps and when deforestation occurs, a fraction of the forest biomass is lost to the atmosphere immediately as CO<sub>2</sub> and a fraction is assumed to be harvested and transferred to three pools wood products of different lifetimes. The establishment of a secondary ecosystem, generally cropland or pasture, produces disequilibrium between NPP driving the input of carbon to the soil pools and heterotrophic respiration, which generates delayed soil CO<sub>2</sub> emissions during several years following deforestation. In the version of ORCHIDEE used to estimate  $E_{LUC}$ , soil carbon stocks of forests being cut in a grid cell a given year are diluted with soil carbon of existing croplands and pastures, so that  $E_{LUC}$  can only be estimated as the simulated NEE difference between a simulation with LUC and one without LUC. This method to estimate  $E_{LUC}$ , used in most global dynamic vegetation models, accounts for a “missed sink capacity” and thus likely overestimates  $E_{LUC}$  as shown by Gasser and Ciais (2013) and Pongratz et al. (2009). Monthly  $E_{LUC}$  at 0.5° resolution from (4.1) was distributed uniformly each week of the same month, and then interpolated to the transport model grid.

$E_{LUC}$  from (4.1) is the sum of deforestation associated to fires and deforestation due to other practices since ORCHIDEE does not distinguish between the two processes. Typically, natural fires within moist Amazonian forests are rare. When they occur, burning areas are limited in size and fire intensity is low due to high humidity of vegetation and litter. Most fires within the Amazon Basin are caused by human activities (Pivello, 2011) associated to LUC, and they occur during the dry season in the southern part of the Amazon basin, mainly. On the other hand, in the cerrado grasslands fire occurrence is linked to both natural processes and human activities. Natural fires in Brazil are caused by lightning, and lightning-associated fires in the cerrado are frequent (Ramos-Neto and Pivello, 2000), whereas anthropogenic fires are part of practices to remove natural vegetation, shifting

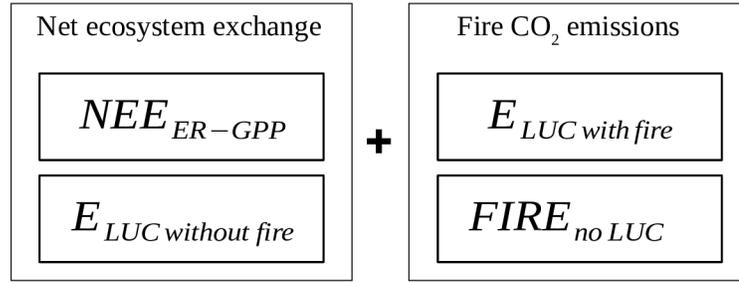
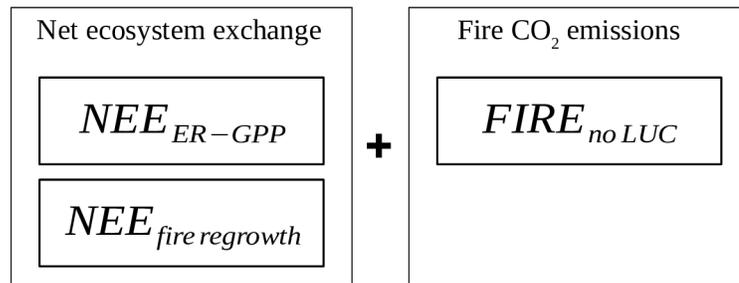
a) Net surface-atmosphere CO<sub>2</sub> exchange on areas subject to LUCb) Net surface-atmosphere CO<sub>2</sub> exchange on areas not subject to LUC

Figure 4.4: Fluxes accounted for in the prior estimates of net ecosystem exchange and fire CO<sub>2</sub> emissions in (a) areas subject to land use change (LUC) and in (b) areas not subject to LUC.

cultivation, and crop management. In the cerrado, vegetation lost to wildfires is usually assumed to recover quickly, so that carbon released is compensated by vegetation regrowth within a short period of few years. Figure 4.4 depicts the fluxes that are accounted for in the prior flux estimates of NEE and fire emissions.

Thus, Figure 4.4a illustrates that in areas subject to LUC, the final NEE product used for the prior estimate and observation operator of the inversion is estimated as

$$NEE = NEE_{ER-GPP} + E_{LUC\ without\ fire} \quad (4.2)$$

where  $NEE_{ER-GPP}$  corresponds to the initial NEE estimate (calculated as  $ER - GPP$ ) from ORCHIDEE.  $E_{LUC\ without\ fire}$  represents deforestation emissions due to human practices other than fires.

Emissions due to deforestation fires ( $E_{LUC\ fire}$ ) and emissions caused by natural fires ( $FIRE_{no\ LUC}$ ) in a given grid-cell were not accounted for in the NEE, but in the fire emissions (Figure 4.4a). On the other hand, in non-LUC affected areas (Figure 4.4b) where wildfires take place, typically in cerrado areas, the NEE is estimated as

$$NEE = NEE_{ER-GPP} + NEE_{fire\ regrowth} \quad (4.3)$$

where  $NEE_{fire\ regrowth}$  represents the ecosystem recovery from natural fire events.  $NEE_{fire\ regrowth}$  is a flux that is assumed to compensate locally for annual emissions from natural fires, with the regrowth uptake being uniformly distributed throughout the year, as

in Chevallier et al. (2010), and set to balance annually fire emissions. The two components of  $E_{LUC}$ ,  $E_{LUC\ fire}$  and  $E_{LUC\ without\ fire}$  were separated using the data layers of deforestation and non-deforestation fires in the Global Fire Emissions Database version 4 (GFED4 Giglio et al., 2013). In one grid-cell:

$$E_{LUC\ without\ fire} = \begin{cases} 0 & \text{if } E_{LUC} \leq E_{LUC\ fire}, \\ E_{LUC} - E_{LUC\ fire} & \text{if } E_{LUC} \geq E_{LUC\ fire} \end{cases} \quad (4.4)$$

The excess of fire emissions  $E_{LUC\ fire} - E_{LUC}$  is attributed to  $FIRE_{no\ LUC}$ .  $E_{LUC\ fire}$  and  $FIRE_{no\ LUC}$  are part of  $E_{FIRE}$  emissions, which are adjusted separately in the inversion. Figure 4.5 shows the distribution of CO<sub>2</sub> emissions from  $E_{LUC}$ ,  $E_{LUC\ without\ fire}$  and  $FIRE_{no\ LUC}$ .

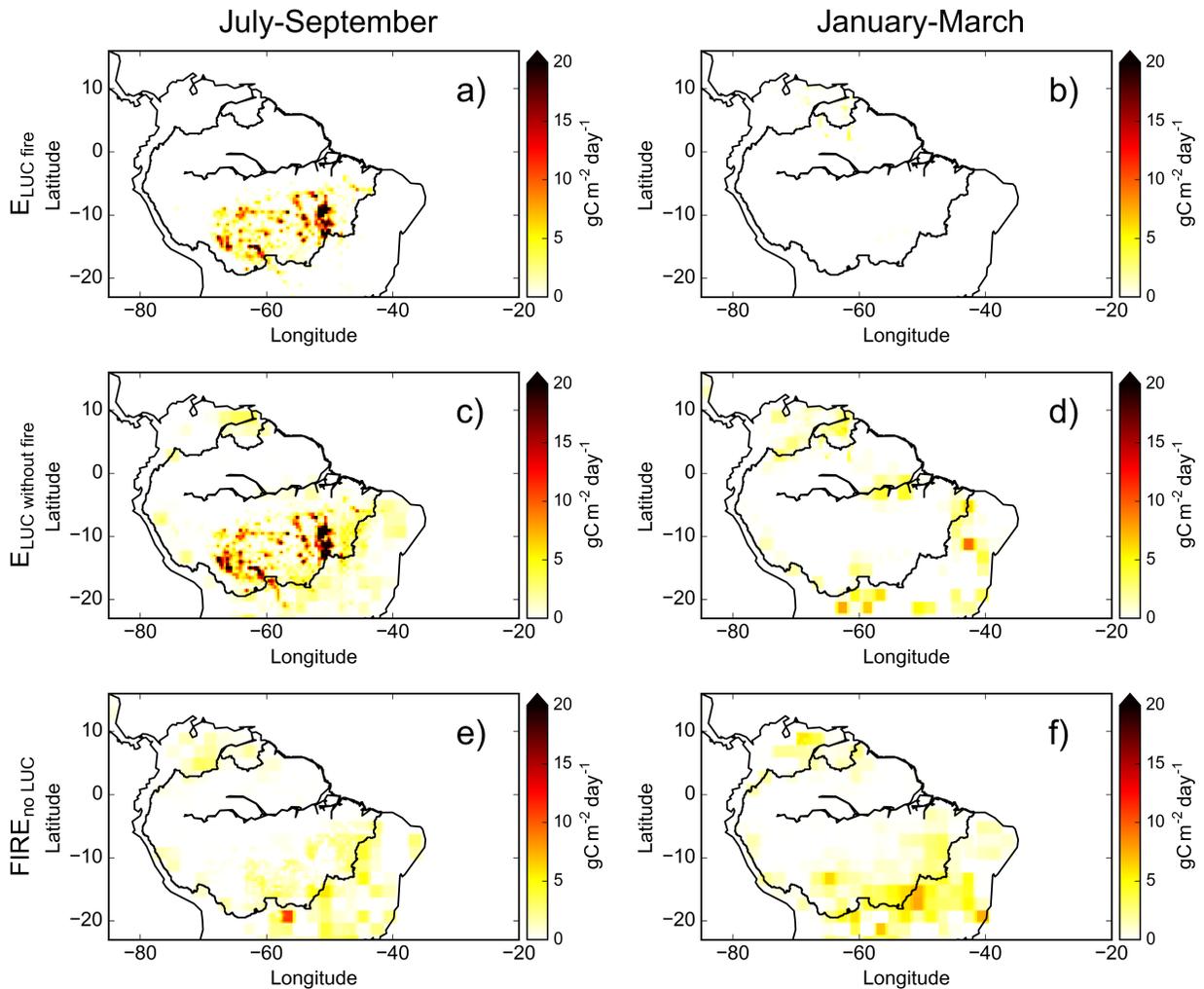


Figure 4.5: Spatial distribution of Amazon CO<sub>2</sub> emissions from (a, b) deforestation fires ( $E_{LUC\ fire}$ ), (c, d) deforestation not due to fire ( $E_{LUC\ without\ fire}$ ), and (e, f) non-deforestation fires ( $FIRE_{no\ LUC}$ ).

### Fire CO<sub>2</sub> Emissions ( $E_{\text{FIRE}}$ )

$E_{\text{FIRE}}$  emissions were taken from the monthly totals of the GFED4 database, originally at  $0.25^\circ$  horizontal resolution. The distribution of CO<sub>2</sub> emissions from deforestation and non-deforestation fires (with deforestation fires being considered a fraction of  $E_{\text{LUC}}$ ) within the simulation domain for 2010, are shown in Figure 4.6. Higher  $E_{\text{FIRE}}$  emissions are located in the south and southeastern borders of the Amazon basin. Two emission epicenters can be identified approximately east and west of  $55^\circ$  W in the figure. To consider the possibility of different fire emissions, e.g. due to different fuel types and fire regimes around these two areas, I split the region in two sub-regions (Figure 4.3b) for the inversion. In the inversion,  $E_{\text{FIRE}}$  emissions were adjusted every week.

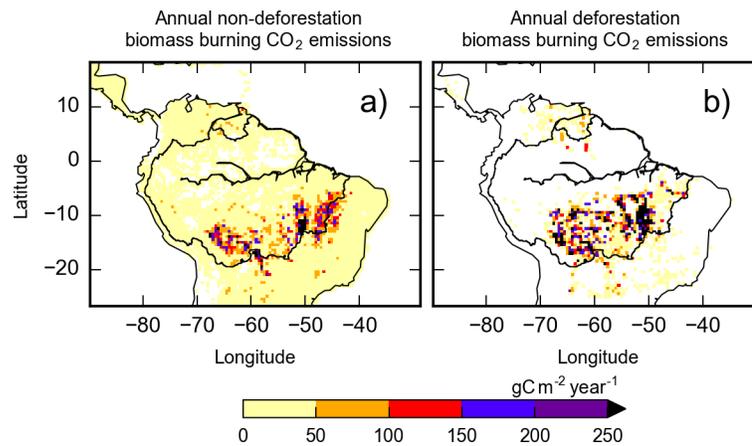


Figure 4.6: Annual biomass burning CO<sub>2</sub> emission from (a) non-deforestation and (b) deforestation fires for 2010, obtained from GFED4, at the native resolution of  $0.25^\circ$ .

### Air-sea fluxes

Despite the results from sensitivity tests in Chapter 3 indicating a very small signature of ocean fluxes on CO<sub>2</sub> concentrations data or gradients within Amazonia, ocean fluxes were optimized in the inversion. As in Chapters 2 and 3, ocean fluxes are taken from the climatology of air-sea CO<sub>2</sub> exchange from Takahashi et al. (2009). Ocean fluxes from the Atlantic and Pacific oceans, within the simulation domain, were optimized separately (Figure 4.3c). In the inversion, ocean fluxes are adjusted every month.

### Fossil fuel CO<sub>2</sub> emissions

To provide a full simulation of the CO<sub>2</sub> concentration in the Amazon basin, an estimate of the CO<sub>2</sub> anthropogenic emissions must be provided to the atmospheric transport model within the observation operator (term  $y_{\text{fix}}$  in equation (4.5); see Section 4.1.4), even though they are not optimized. These anthropogenic CO<sub>2</sub> emissions were prescribed with the monthly totals of the Emission Database for Global Atmospheric Research 2010 Fast-Track version 4.2 (EU-JRC/PBL, 2013), scaled with annual global totals from the Global Carbon Project 2013 (Boden et al., 2013).

### 4.1.3 Prior uncertainty covariance structure

The prior uncertainty in the control parameters, in this case the scaling factors for sub-regional weekly/monthly fluxes, and their spatial and temporal correlation are represented in the prior error covariance matrix  $\mathbf{B}$ . Prior uncertainties of NEE and ocean fluxes were adjusted to match the monthly uncertainty ( $1-\sigma$ ) of the respective fluxes within the simulation domain, based on the setup of prior uncertainties for the global atmospheric inversion of Chapter 2, and on the description of error correlations detailed below. Prior uncertainties of each NEE sub-region/week were in the range of  $\sim 0.2 - 8 \text{ gC m}^{-2} \text{ day}^{-1}$ , and  $\sim 0.02 - 0.06 \text{ gC m}^{-2} \text{ day}^{-1}$  for each sub-region/month for ocean fluxes. For  $E_{\text{FIRE}}$  emissions, I calculated the prior uncertainty as the standard deviation of the monthly CO<sub>2</sub> emission in the simulation domain based on three datasets: GFED4, the Global Fire Assimilation System (GFASv2.1) (Kaiser et al., 2012), the Fire Inventory of the National Center for Atmospheric Research (FINN) (Wiedinmyer et al., 2011). All three datasets are based on satellite observations, but differ in the input data. Both GFED4 and FINN calculate emissions based on burned area. But while GFED4 combines satellites observation of burned area and active fire data (i.e. fire counts), FINN uses only active fires. Emissions from GFASv2.1, on the other hand, are based on fire radiative power. Prior uncertainty for  $E_{\text{FIRE}}$  emissions for each region/week were in the range of  $\sim 0.14 - 0.92 \text{ gC m}^{-2} \text{ day}^{-1}$ .

The spatial correlation of the uncertainties between the fluxes in two sub-regions, was estimated with the relationship  $r = e^{-d/L}$ , where  $r$  is the error correlation,  $d$  is a distance (in kilometers) separating the geometrical centers (centroids) of the regions, and  $L$  is a scale length. For NEE and ocean fluxes,  $L_{\text{NEE}}$  and  $L_{\text{ocean}}$  were set to 500 km and 1000 km, respectively. These correlations length scales are similar to the values used in Chapter 2, even though the spatial resolution of the control vector is different, which may imply different correlation length scales when considering uncertainties at sub-regional scale. For  $E_{\text{FIRE}}$  emissions, the chosen error correlation length scale ( $L_{\text{fire}}$ ) is more arbitrary, and I used 500 km. In the future, comparing the structure of the three fire emission datasets could allow a better formulation for. Spatial error correlation lengths of between 500 km (e.g. Yin et al., 2016) and 1000 km (e.g. Hooghiemstra et al., 2011) have been used in previous CO inversion studies, yet for different control vector resolutions. For all emission types, temporal error correlations were modeled following an exponential decay function with a time scale of one month, as in Chapter 2, in which the temporal resolution of the control vector was similar. The a priori uncertainties between different emission types are assumed to be uncorrelated.

### 4.1.4 Observation operator

Fluxes are projected into the concentration space through the linear observation operator:

$$H : \mathbf{x} \mapsto \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{y}_{\text{fix}} \quad (4.5)$$

Section 1.1.3 introduced the fact that  $\mathbf{H}$  can be seen as the combination of three operators:  $\mathbf{H} = \mathbf{H}_{\text{sample}}\mathbf{H}_{\text{tran}}\mathbf{H}_{\text{dist}}$ .  $\mathbf{H}_{\text{dist}}$  takes the grid-based fields of NEE,  $E_{\text{FIRE}}$  emissions and ocean fluxes to predefine spatial and temporal flux patterns within control sub-regions and temporal windows. The operator applies each scaling factor corresponding to the control vector to rescale the spatial and temporal distribution of fluxes on the grid of the transport model in each sub-region, for each week or month. The operator  $\mathbf{H}_{\text{tran}}$  is the transport model. I simulated CO<sub>2</sub> transport with the CHIMERE model at a resolution of

~35 km, driven with the fields from BRAMS and ECMWF (Chapter 3). Finally,  $\mathbf{H}_{\text{sample}}$  samples and process the simulated CO<sub>2</sub> fields to derive a vector of simulated observations corresponding to the selected observation vector (Section 4.1.5). The simulated 3D fields were sampled by selecting the transport model’s grid-cell whose center was closest to the position (horizontal and vertical) of each individual measurement from GA2014, and then, depending on the observation vector chosen (see below), the resulting simulated data set was potentially processed (filtered and combined) into horizontal gradients along the wind direction close to the surface, for different vertical bins, or for the total columns along the measurement profiles.

$\mathbf{y}_{\text{fix}}$  contains the signatures of anthropogenic CO<sub>2</sub> emissions and of the model CO<sub>2</sub> boundary conditions on the atmospheric CO<sub>2</sub> concentration. Those signatures are “fixed” in the sense that they are not optimized by the inversion, but must be accounted for when comparing simulations and CO<sub>2</sub> measurements. This signature of anthropogenic CO<sub>2</sub> emissions and CO<sub>2</sub> boundary conditions was simulated with the same transport model configuration as that used to build  $\mathbf{H}$  for a given inversion configuration, based on the estimate of the anthropogenic CO<sub>2</sub> emissions described in Section 4.1.2, and on boundary conditions from the post-inversion CO<sub>2</sub> fields from the dedicated inversion MACC+ used in Chapter 3 (section 3.1.1).

#### 4.1.5 Observation vector and error covariance structure

CO<sub>2</sub> fluxes were optimized by assimilating four different observation vectors whose general principle was described in Chapter 3. In this chapter I used a refined definition of these observation vectors, in particular regarding the selection of the horizontal gradients as a function of the wind direction, rather than considering fixed gradient configurations based on the mean atmospheric circulation in the Amazon basin. I assimilated the individual measurements in vertical profiles and three types of horizontal gradients between profile sites along the dominant wind direction: gradients within 1 km vertical bins between 0 and 5 km a.s.l., gradients between 0 and 1 km a.s.l., and the integral of the horizontal gradients from 0 to 5 km a.s.l. The four observation vectors build on the measurements of GA2014. The observation vector and the number of observations assimilated in each period are summarized in Table 4.2. The following provides more details on the definition of the observations vectors, on the resulting datasets, and on the estimate of observation errors for each observation vector. A general assumption is that the observation error is dominated by transport model errors, and that there is no correlation, either spatial or temporal, of the observation error between two different observations, so that  $\mathbf{R}$  is systematically set up diagonal. The 2-week frequency of the measured profiles at a given site supports not accounting for temporal correlations, while the distance between the measurement sites (500 to 1700 km) supports not accounting for horizontal correlations.

However, the assumption of negligible vertical correlations of the transport errors between individual measurements or horizontal gradients may be strong, given the vertical patterns of differences between CH-BRAMS and CH-ECMWF in Chapter 3. While the weight of the transport errors might thus be underestimated in the setup of  $\mathbf{R}$ , carrying out inversions with both transport models provides a robust indication about the impact of the transport error on the inverted fluxes.

Table 4.2: Observation vectors assimilated.

Identifier	Description	Number of observations	
		July–September	January–March
PROFILE	Individual CO <sub>2</sub> measurements at each site	257	280
HGRAD	CO <sub>2</sub> gradient at different altitudes (1000 m bins)	10	20
HSURF	CO <sub>2</sub> gradient close to the surface (0 to 1000 m)	10	20
HINTEG	Vertical average of CO <sub>2</sub> gradient (0 to 4000 m)	10	20

### Assimilation of individual CO<sub>2</sub> measurements at each sampling site

In the inversion experiments called PROFILE, each individual CO<sub>2</sub> measurement of the profiles from GA2014 was assimilated, i.e. 257 observations during JAS, and 280 during JFM were assimilated from four aircraft sampling sites. The observation error was derived for each site (i.e. different error estimates for different sites) as the standard deviation of the difference between two the transport simulations CH-BRAMS and CH-ECMWF forced with the following components (Chapter 3): the NEE from ORCHIDEE (not including land use change or fire CO<sub>2</sub> emissions), E<sub>FIRE</sub> emissions from GFED3 database, ocean fluxes from Takahashi et al. (2009), fossil fuel CO<sub>2</sub> emissions from EDGAR4.2FT and post-inversion CO<sub>2</sub> fields from MACC+ as boundary conditions. Chapter 3 showed that modeled CO<sub>2</sub> differences from the two transport fields were larger within 1 – 2 km than at higher altitudes. Therefore I derived two typical estimates of the transport error, one within and one above the planetary boundary layer (PBL), based on statistics of the differences between the two models at the measurements’ time and locations in these 2 zones of the atmosphere. I chose a typical PBL height of 1000 m based on the estimate of Fisch et al. (2004) at a forest site in southwestern Amazonia during dry or wet seasons. The observation error ranged between ~2 and 5 ppm.

### Assimilation of CO<sub>2</sub> horizontal gradients at different altitudes between sites

In this experiment called HGRAD, CO<sub>2</sub> gradients between profile sites along the dominant wind direction (i.e., downwind minus upwind) were assimilated. Due to differences in time and altitude where observations were collected, observations were averaged in vertical bins every 1000 m for each profile. Ideally, in order to maximize the potential for decreasing the signature of fluxes upwind the “upwind profile” and enhance the signature of the NEE between the two vertical profiles, the time interval between two profile sites should be the air-mass travel time between them (Staufer et al., 2016). But the measured profiles do not fully satisfy this condition. Therefore, to retain more data, gradients were calculated considering that profiles at the downwind site could be up to 5 days later than the upwind site. To account for those situations when the average wind deviated from the direction between sites and retain more data, I considered a broader range of wind directions (Boon et al., 2016; Bréon et al., 2015). Wind direction was based on the meteorology simulated with BRAMS or ECMWF. I considered that horizontal gradients were suitable for enhancing

the signal of NEE—and reducing the influence of uncertainties in other types of fluxes—if the angle between the direction from the upwind to the downwind site and the wind direction at the downwind site was  $\pm 45^\circ$ . The values of ranged between  $\sim 0.7$  and 1.6 ppm.

### Assimilation of CO<sub>2</sub> horizontal gradients within the PBL between sites

This experiment is called HSURF and assimilated only the horizontal gradient of CO<sub>2</sub> in the PBL, i.e. from 0 to 1000 m.a.s.l. Vertical profiles and horizontal CO<sub>2</sub> gradients in Chapter 3 showed a strong sensitivity to NEE and E<sub>FIRE</sub> emissions in the first kilometer from the surface. Thus, horizontal gradients closest to the surface could yield more information about the NEE, even though they are also strongly influenced by E<sub>FIRE</sub> emissions. The values of were about 1.5 ppm.

### Assimilation of the vertically integrated horizontal gradients of CO<sub>2</sub> between sites

In this experiment called HINTEG, I assimilated the integral of the horizontal gradients, vertically averaged between 0 and 5 km a.s.l. Chapter 3 showed that this type of observation reduced the weight of the uncertainty of boundary conditions relative to that of NEE more than in the other two types of gradient. GA2014 and Gatti et al. (2010) used this type of observation in their mass balance approach. However, GA2014 and Gatti et al. (2010) calculated gradients between aircraft sampling sites in the Amazon basin and surface observations at ocean sites, which defined boundary condition CO<sub>2</sub> concentration for air entering the Amazon basin. I chose not to assume that surface CO<sub>2</sub> observations at the ocean sites could be extrapolated as a homogeneous vertical profile. Instead, I used the gradients between different aircraft sites in the basin, which should leave the fluxes in areas east of the easternmost site not well constrained by the observations. The values of were  $\sim 1$  ppm.

## 4.2 Results: Fit to observed CO<sub>2</sub> data

In this section, I present of the fit of the assimilation to the observed CO<sub>2</sub> concentrations, for the different observation vectors (Section 4.1.5). The terms INV-BRAMS and INV-ECMWF refer to inversions using the two different transport fields presented in Chapter 3.

### PROFILE experiment

The prior and optimized CO<sub>2</sub> profiles are shown in Figures 4.7 and 4.8 for the dry period, and Figures 4.9 and 4.10 for the wet period. In general, after the flux optimization, the distance to observations was effectively reduced compared to the prior. Posterior profiles reproduced reasonably well the shape of the biweekly profiles at each site. In most of cases, the model-observation misfits were within the observation error prescribed in the inversion (1.4 to 4.5 ppm; Section 4.1.5), except when the observed profiles exhibited more vertical structure than the posterior profiles, e.g., around the points of sharp concentration changes. Large CO<sub>2</sub> adjustments resulted from the flux optimization both in INV-BRAMS and INV-ECMWF. At TAB (Figure 4.8h), corrections close to the surface were as large as 15 ppm in INV-BRAMS and 7 ppm for INV-ECMWF during the dry period, and up to  $\sim 7$  ppm for INV-BRAMS (Figure 4.10f) and  $\sim 13$  ppm for INV-ECMWF (Figure 4.10e) in the wet period. The largest corrections were found in the first kilometer from the surface, i.e. in

the PBL, where CO<sub>2</sub> is most influenced by surface fluxes. Still, in both transport models corrections of up to  $\sim 6$  ppm were found above 2 km both in the dry and wet periods. In the PBL, a better fit to the observations would be expected than in altitude. However, during both dry and wet periods, a better fit to observations was found in altitude rather than in the PBL (e.g. Figures 4.7a, Figure 4.9l, Figure 4.10(a – d)).

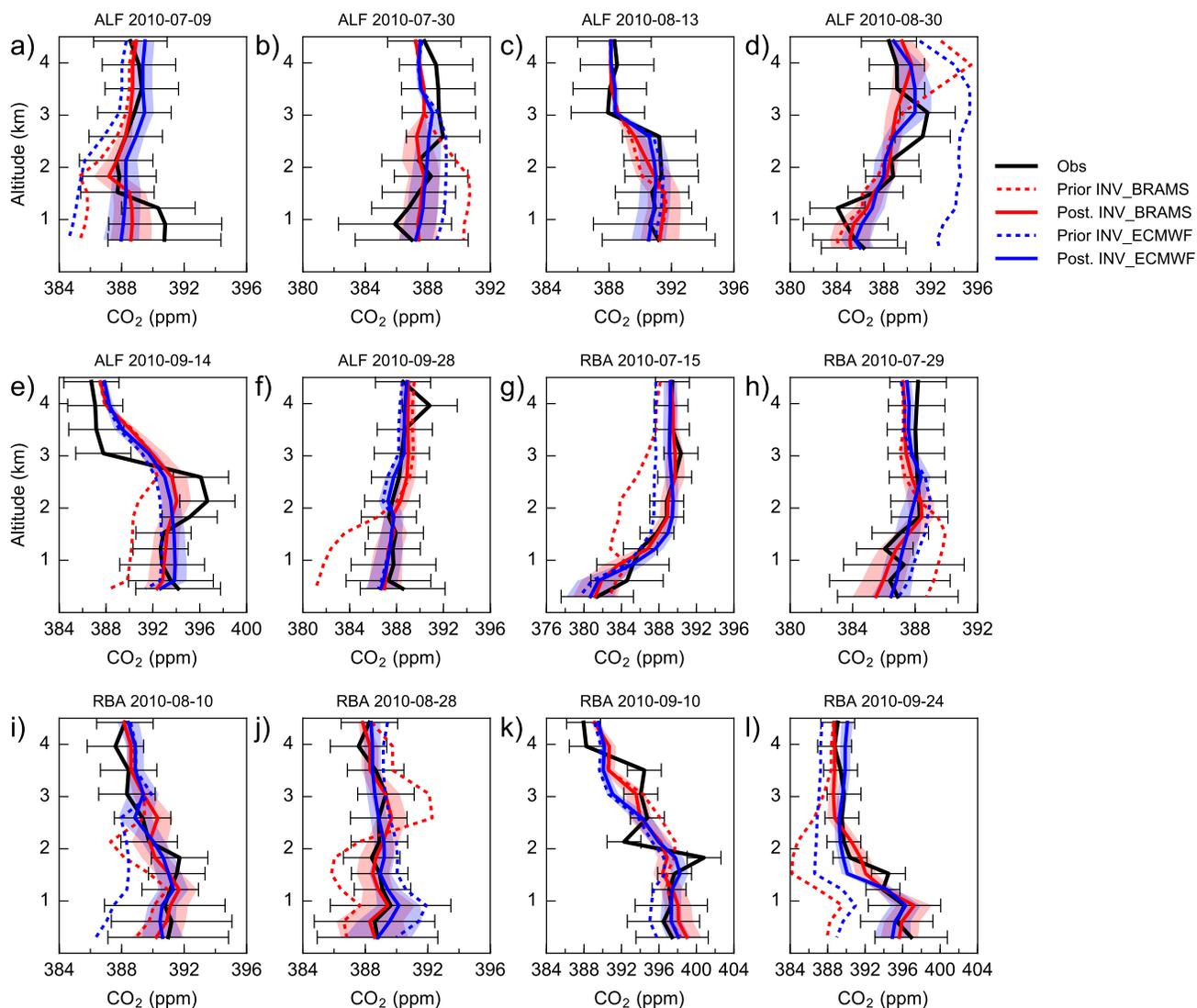


Figure 4.7: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> vertical profiles for the dry period of 2010 (July – September). Vertical profiles were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior profiles depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

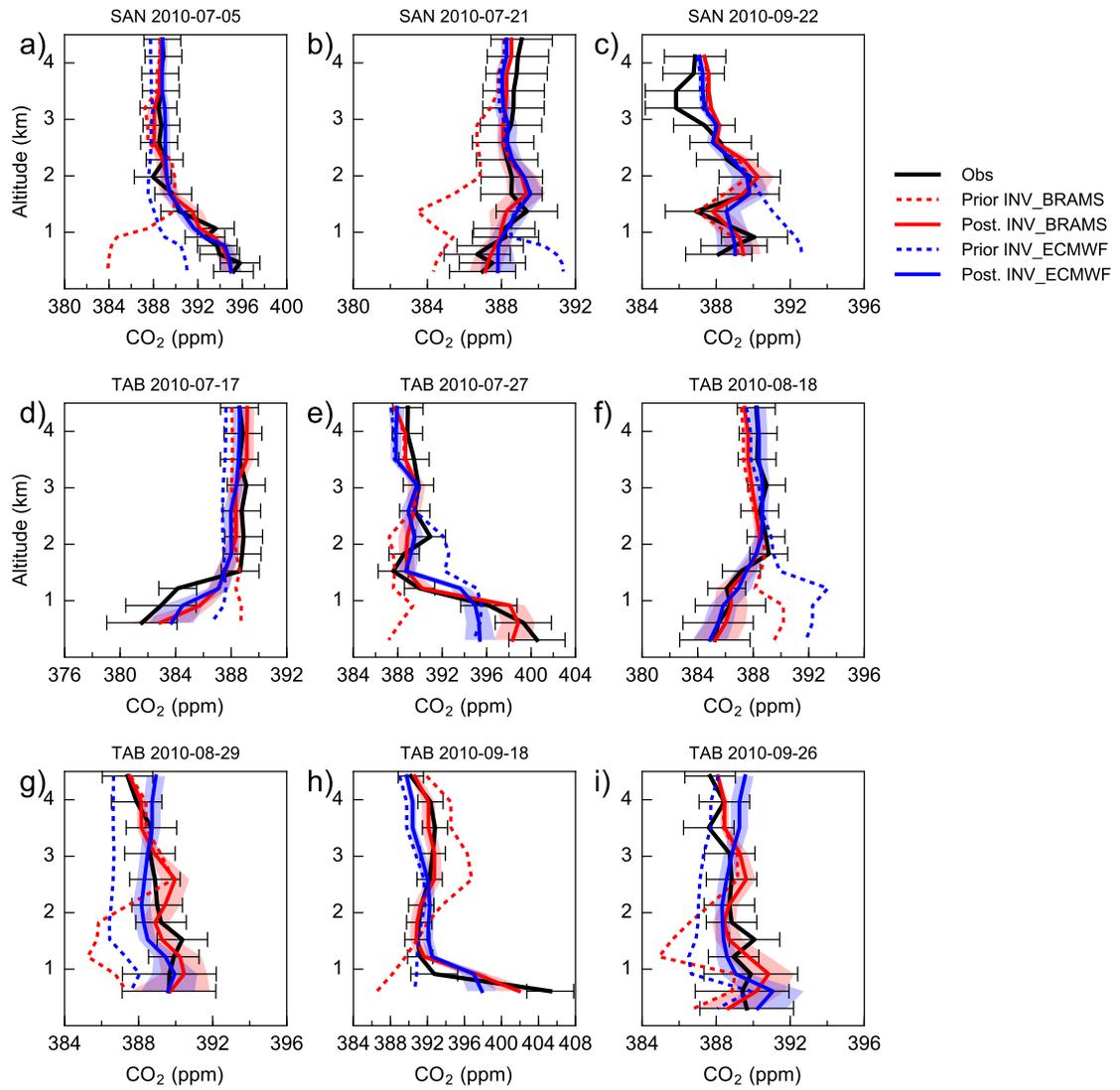


Figure 4.8: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> vertical profiles for the dry period of 2010 (July – September). Vertical profiles were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior profiles depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

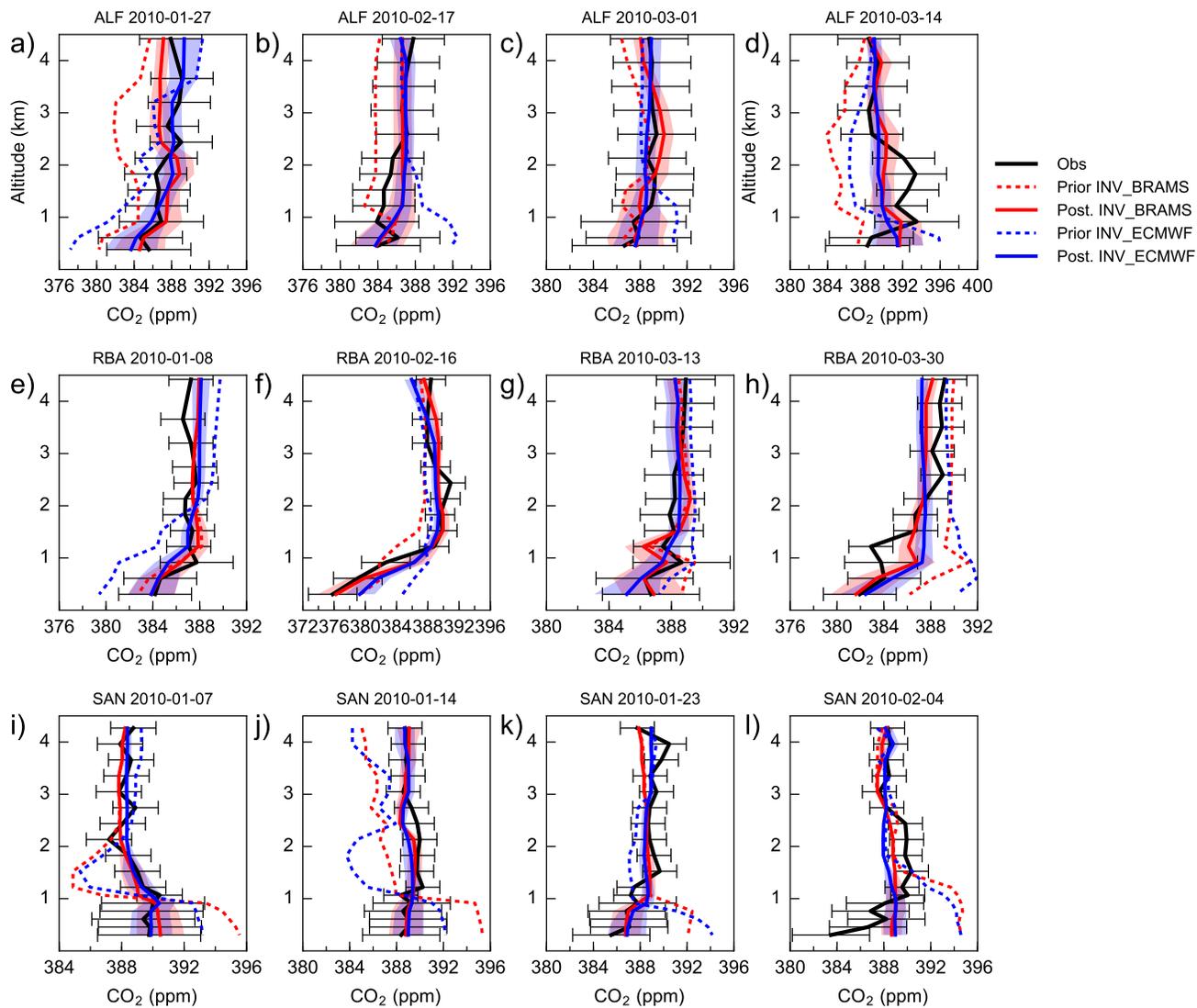


Figure 4.9: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> vertical profiles for the wet period of 2010 (January – March). Vertical profiles were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior profiles depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

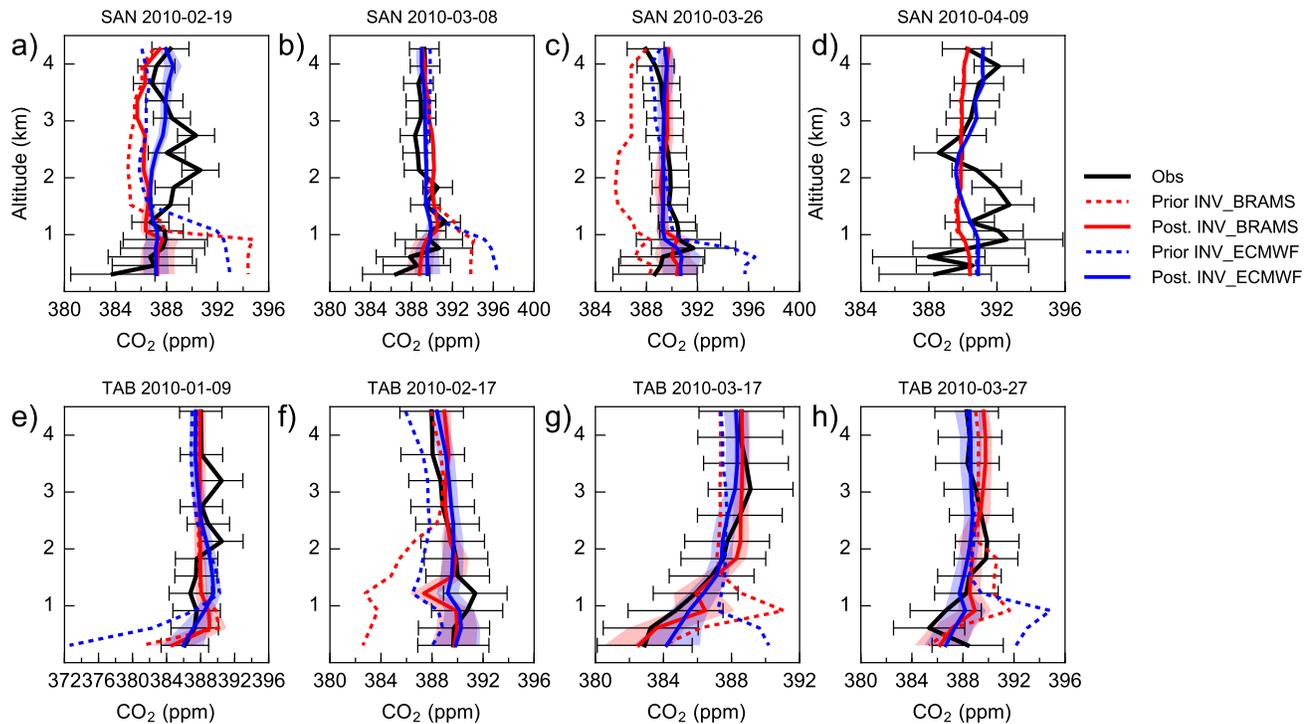


Figure 4.10: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> vertical profiles for the wet period of 2010 (January – March). Vertical profiles were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior profiles depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

The results of Chapter 3 showed that CO<sub>2</sub> in altitude was also sensitive to the choice of different NEE, given the coupling between vertical and horizontal transport in the Amazon, so that important corrections to the profiles in altitude caused by NEE optimization remain plausible. The residual misfits in the PBL not corrected after the inversion could reflect errors in the transport models for PBL mixing, given that ECMWF has a simple land surface scheme, and that BRAMS with JULES may also have biases in the simulation of surface energy budget, which impacts the timing of PBL mixing and the mixed layer height. Another plausible explanation is the fact that my source NEE field from ORCHIDEE was at coarse resolution (2°) and only optimized over the large sub-regions of Figure 4.3a., i.e. potential impact of aggregation errors, even though the choice of the control sub-region was made to minimize them.

Despite lower scores in the PBL in some profiles, optimized CO<sub>2</sub> showed a good statistical fit to the observations when considering all profiles and all altitudes together. Prior root-mean-square error (RMSE) across all sites and altitudes decreased from ~3 ppm to ~1 ppm in posterior concentrations for both INV-BRAMS and INV-ECMWF during both periods. Desroziers et al. (2005) proposed a consistency diagnostic of the analysis-error statistics in the observation space in order to evaluate, and potentially refine, the configuration of the prior uncertainties and observation errors in the inversion setup. Such

diagnostics were applied to NEE atmospheric inversions by [Chevallier and O'Dell \(2013\)](#). [Desroziers et al. \(2005\)](#) formulated that the variance of observed-minus-a prior (O–B) CO<sub>2</sub> concentration differences  $\mathbf{d}_b^o = \mathbf{y}_o - H(\mathbf{x}_b)$  should be equal to:

$$E[\mathbf{d}_b^o(\mathbf{d}_b^o)^T] = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T \quad (4.6)$$

And the variance of observed-minus-analyzed (O–A) CO<sub>2</sub> concentration difference  $\mathbf{d}_a^o = \mathbf{y}_o - H(\mathbf{x}_a)$  should be equal to:

$$E[\mathbf{d}_a^o(\mathbf{d}_a^o)^T] = \mathbf{R} + \mathbf{H}\mathbf{A}\mathbf{H}^T \quad (4.7)$$

The equalities in (4.6) and (4.7) are satisfied only if the covariance of prior and observational errors,  $\mathbf{B}$  and  $\mathbf{R}$  respectively, reflects the statistics of the actual prior and observation errors. [Desroziers et al. \(2005\)](#) proposed to compute:

$$\left(\widetilde{\sigma}^b\right)^2 = (\mathbf{d}_b^o)^T(\mathbf{d}_b^o)/n = \sum_{j=1}^n (y_j^b - y_j^o)^2/n \quad (4.8)$$

$$\left(\widetilde{\sigma}^a\right)^2 = (\mathbf{d}_a^o)^T(\mathbf{d}_a^o)/n = \sum_{j=1}^n (y_j^a - y_j^o)^2/n \quad (4.9)$$

(4.8) and (4.9) correspond to the mean square (MS) values of the O–B and O–A departures, respectively, to be confronted to the mean of the diagonal elements of the right-hand side of equations (4.6) and (4.7). In Figure 4.11, MS values of the O–A and O–B departures are represented with black dots, and the components of the right-hand side of equations (4.6) and (4.7) as stacked bars. For both INV-BRAMS and INV-ECMWF, the variance equality (4.6) was not satisfied in the case of the prior. Before inversion, the prior error dominated the modeled error balance in both models, and it was larger by itself than the MS of the O–B departures, suggesting that prior errors were overestimated.

After the inversion, it was the observation error that dominated the posterior error budget.  $\mathbf{R}$  values were larger by themselves than the MS of O–A departures, which could suggest that observation error was also overestimated. But such a result could also arise from an over-fit of the data due to a very strong overestimation of  $\mathbf{B}$ . From this diagnostic alone, therefore, it is difficult to say whether errors in both  $\mathbf{R}$  and  $\mathbf{B}$  were overestimated or if only  $\mathbf{B}$  was overestimated, and even in the former case, it is difficult to conclude which type of error was more overestimated (and thus whether data were over-fitted).

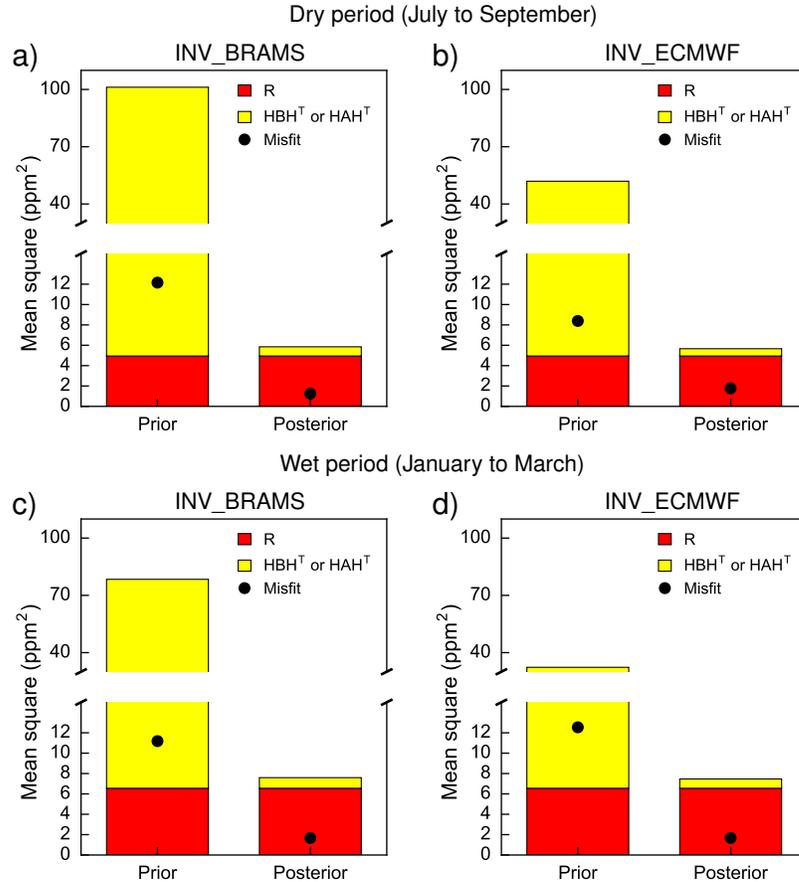


Figure 4.11: Components of Equations 4.6 and 4.7 expressed as mean square (MS) values for the PROFILE inversion. Left-hand terms represented as black dots and right-hand terms as stacked bars.

### HGRAD experiment

Model-observation misfits after the inversion were within the observation error (0.7 – 1.6 ppm; Section 4.1.4) during both dry and wet periods. The vertical structure of the gradients during both periods was adequately fitted by the inversion for both transport models. Corrections were largest in the first layer near the surface (by up to ~12 ppm from prior CO<sub>2</sub>, e.g. Figure 4.13c), but they were also significant in altitude. During the dry period, prior CO<sub>2</sub> gradients in both models increased (became more positive) or decrease (became more negative) towards the surface, which was opposite to the observations Figure 4.12(a – c). This is also seen in the wet period (Figure 4.13(a, c and e)). The inversion rectified those misfits. In general, during both periods, posterior gradients simulated with both models showed a good fit to the observed gradients from the surface up to 5 km.

Figure 4.14 shows the components of equations (4.6) and (4.7). During both dry and wet periods for both models, before the inversion the prior uncertainty dominated the simulated error budget and, again, the diagnostics indicated that prior uncertainties were likely overestimated since they were larger by themselves than the MS of the O–B misfits.

INV-ECMWF was closer to satisfying (4.6) and (4.7) than INV-BRAMS, which indicates a strong effect of the atmospheric transport on the projection of the prior uncertainties in the observation space and questions the assumption that  $\mathbf{B}$  was overestimated. But when analyzing the results for INV-BRAMS the diagnostics of Desroziers et al. (2005) strongly supports that  $\mathbf{B}$  was overestimated. Therefore, these findings question the robustness of such diagnostics, which strongly rely on the assumptions of the inversion theory and on the assumptions I made for the structure of the correlations in  $\mathbf{B}$  and  $\mathbf{R}$ . After the inversion, observation error dominated the modeled error budget. While this result supports the overestimation of  $\mathbf{R}$ , it could also be a consequence of the overestimation of  $\mathbf{B}$ . Therefore, from these diagnostics alone it is difficult to conclude whether  $\mathbf{B}$  or  $\mathbf{R}$ , or both, were overestimated, and in the latter case, in which one the overestimation was more problematic.

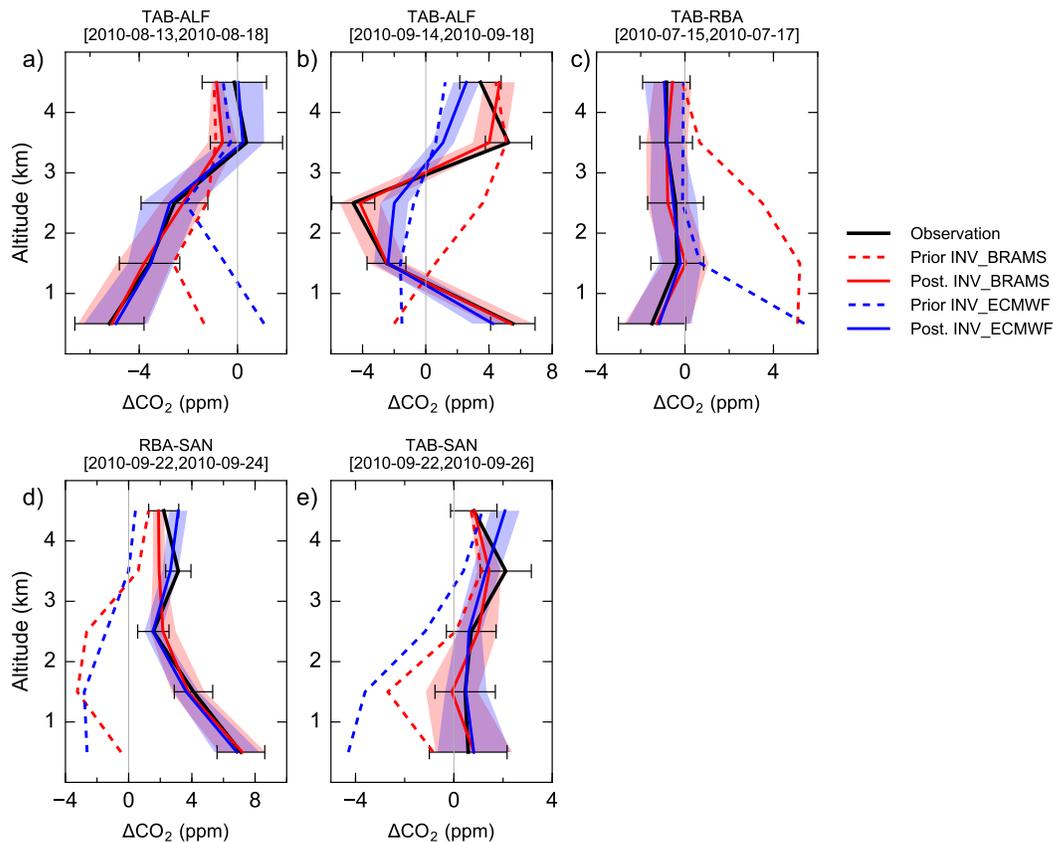


Figure 4.12: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> horizontal gradient for the dry period of 2010 (July – September). Gradients depict the difference of the mean CO<sub>2</sub> for 1000-m layers from 0 to 5 km between downwind and upwind sites. Gradients were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior curves depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

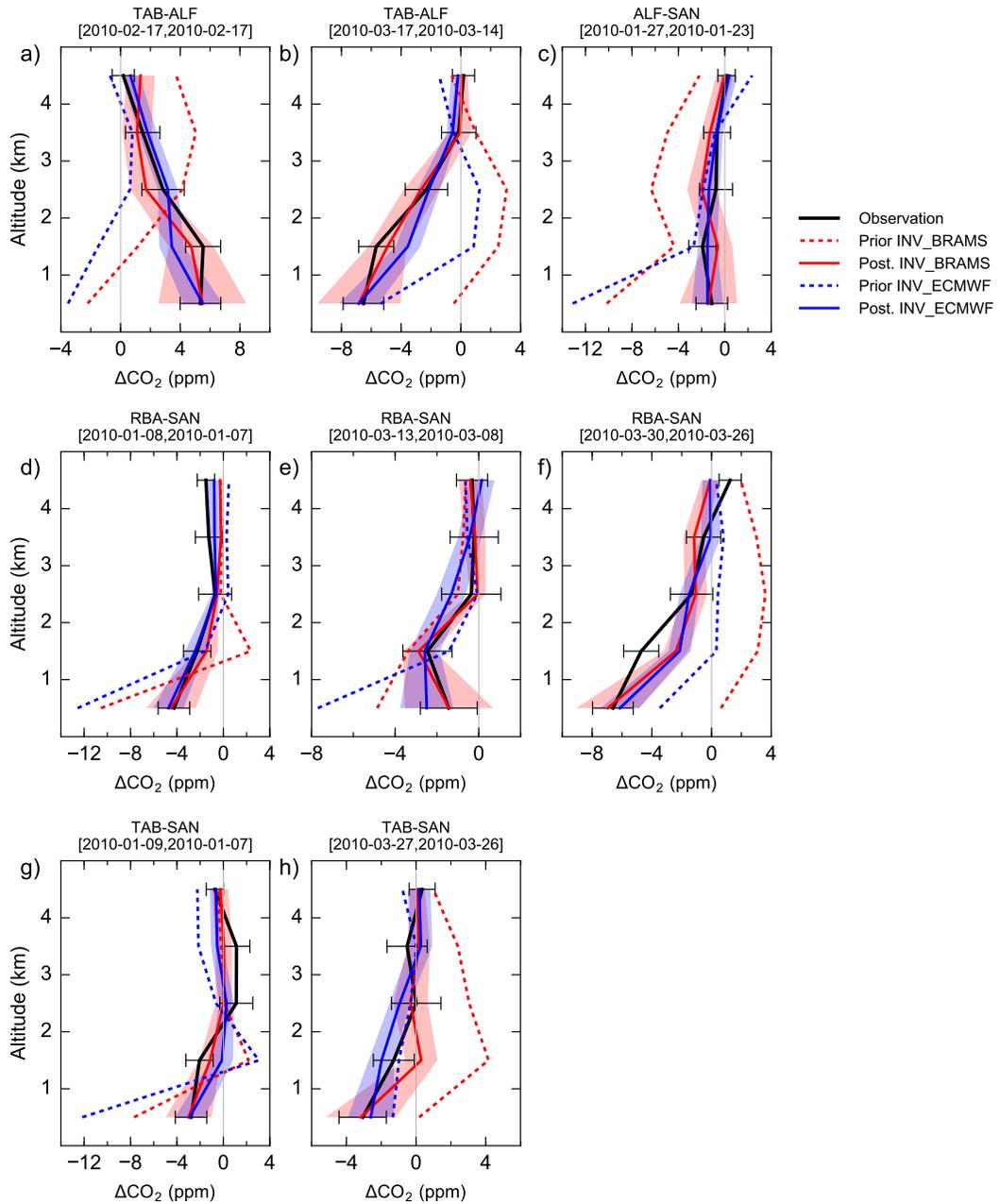


Figure 4.13: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> horizontal gradients for the wet period of 2010 (January–March). Gradients depict the difference of the mean CO<sub>2</sub> for 1000-m layers from 0 to 5 km between downwind and upwind sites. Gradients were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Shaded areas on the posterior curves depict the posterior uncertainty for both model estimates. Error bars represent observation error, estimated as the standard deviation of the between-model misfit.

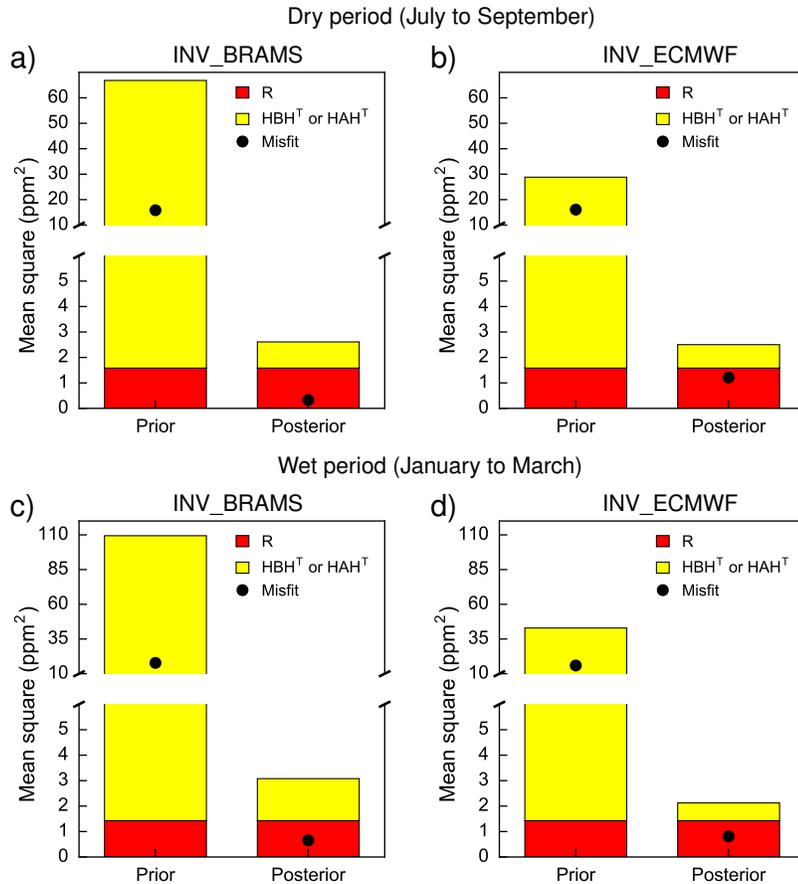
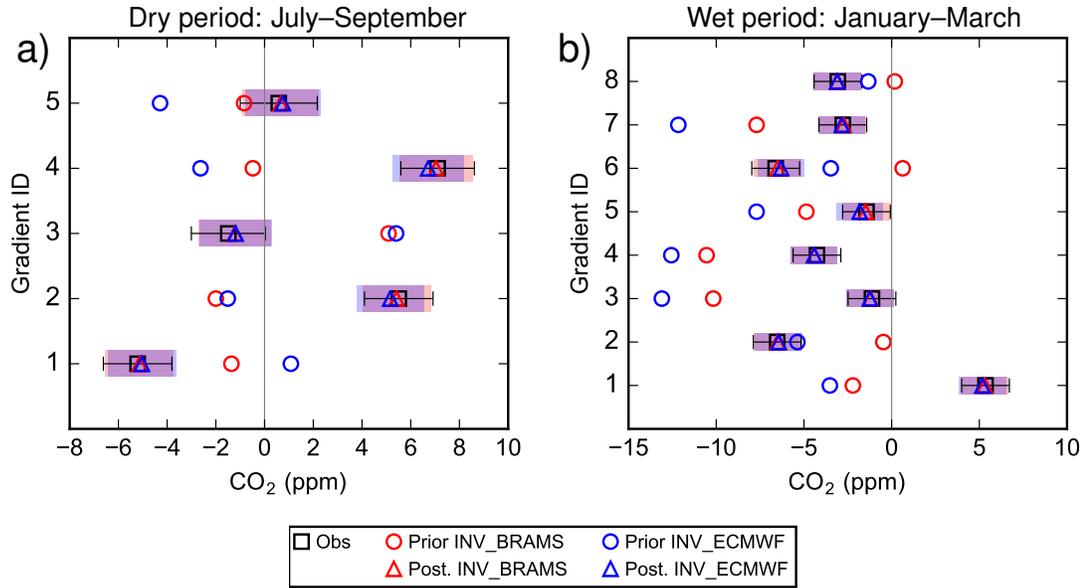


Figure 4.14: Components of Equations 4.6 and 4.7 expressed as mean square (MS) values for the HGRAD inversion. Left-hand terms represented as black dots and right-hand terms as stacked bars.

### HSURF experiment

For both models during the dry or wet periods, large prior misfits ranging from 1.2 to 12 ppm were reduced to 0 – 0.4 ppm (Figure 4.15). The diagnostics of Desroziers et al. (2005) indicated that during both dry and wet periods the uncertainty budget before the inversion was dominated by the prior uncertainty (Figure 4.16), and that prior uncertainties were overestimated (the right-hand side of (4.6) of is larger than the MS of the O–B misfits). After the inversion, the prior and observation errors had a similar weight in the posterior error budget (Figure 4.16), so that with this diagnostic alone it is not possible to conclude in which one, or, the overestimation was more problematic.



ID	Transect	Date upwind	Date downwind
5	TAB-ALF	2010-09-22	2010-09-26
4	TAB-ALF	2010-09-22	2010-09-24
3	TAB-RBA	2010-07-15	2010-07-17
2	RBA-SAN	2010-09-14	2010-09-18
1	TAB-SAN	2010-08-13	2010-08-18

ID	Transect	Date upwind	Date downwind
8	TAB-ALF	2010-03-26	2010-03-27
7	TAB-ALF	2010-01-07	2010-01-09
6	ALF-SAN	2010-03-26	2010-03-30
5	RBA-SAN	2010-03-08	2010-03-13
4	RBA-SAN	2010-01-07	2010-01-08
3	RBA-SAN	2010-01-23	2010-01-27
2	TAB-SAN	2010-03-14	2010-03-17
1	TAB-SAN	2010-02-17	2010-02-17

Figure 4.15: Observed (**black**) and simulated prior and posterior surface (0 to 1000 m above sea level) CO<sub>2</sub> horizontal gradient for **(a)** dry (July – September) and **(b)** wet (January – March) periods, between downwind and upwind sites. Gradients were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Numbers in the y-axis correspond to the observations listed in the respective tables below the plots. Shaded areas on the posterior curves depict the posterior uncertainty for both model estimates. Error bars represent the observation error.

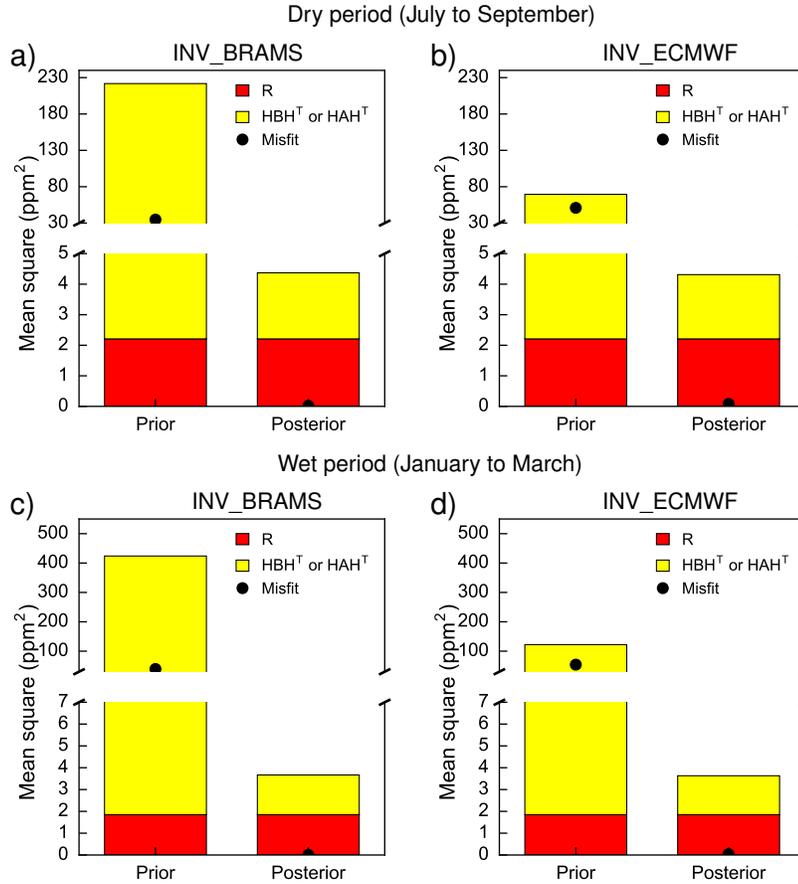
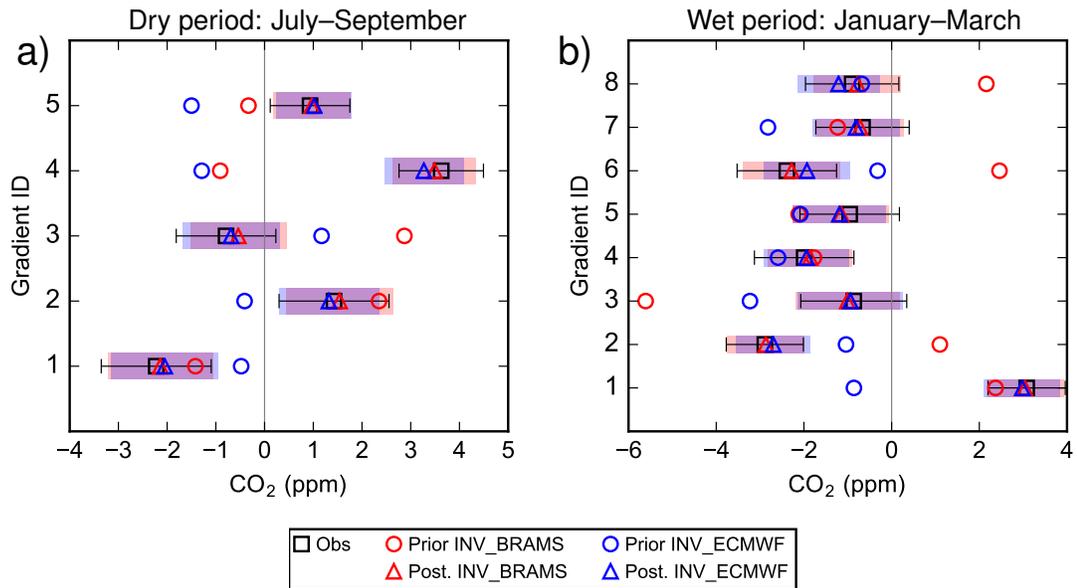


Figure 4.16: Components of Equations 4.6 and 4.7 expressed as mean square (MS) values for the HSURF inversion. Left-hand terms represented as black dots and right-hand terms as stacked bars.

### HINTEG experiment

The horizontal gradients integrated along the vertical were well reproduced in the posterior gradients (Figure 4.17). Considering both inversion models and both dry and wet periods altogether, prior departures of  $\sim 0.2$  to  $4.9$  ppm were reduced to  $\sim 0.2 - 0.5$  ppm. During the dry period, the prior root mean square error (RMSE) of  $2.7$  and  $1.7$  ppm in INV-BRAMS and INV-ECMWF, respectively, was reduced to  $0.1$  and  $0.2$  ppm after the inversion. Similar performance was achieved during the wet period, where prior RMSE of  $\sim 3$  and  $2$  ppm in INV-BRAMS and INV-ECMWF, respectively were reduced to  $\sim 0.2$  ppm in both cases.

The consistency diagnostics proposed by Desroziers et al. (2005), showed that prior uncertainties prescribed in  $\mathbf{B}$  dominated the error balance before the inversion, and suggest that  $\mathbf{B}$  was overestimated. After the inversion, like in the HSURF experiments, the posterior uncertainty was as important in the posterior error budget as the observation errors (Figure 4.18). With this diagnostic alone, though, it is not possible to conclude in which one,  $\mathbf{R}$  or  $\mathbf{B}$ , the overestimation was more problematic.



ID	Transect	Date upwind	Date downwind
5	TAB-ALF	2010-09-22	2010-09-26
4	TAB-ALF	2010-09-22	2010-09-24
3	TAB-RBA	2010-07-15	2010-07-17
2	RBA-SAN	2010-09-14	2010-09-18
1	TAB-SAN	2010-08-13	2010-08-18

ID	Transect	Date upwind	Date downwind
8	TAB-ALF	2010-03-26	2010-03-27
7	TAB-ALF	2010-01-07	2010-01-09
6	ALF-SAN	2010-03-26	2010-03-30
5	RBA-SAN	2010-03-08	2010-03-13
4	RBA-SAN	2010-01-07	2010-01-08
3	RBA-SAN	2010-01-23	2010-01-27
2	TAB-SAN	2010-03-14	2010-03-17
1	TAB-SAN	2010-02-17	2010-02-17

Figure 4.17: Observed (**black**) and simulated prior and posterior CO<sub>2</sub> horizontal gradient, integrated from 0 to 5000 m above sea level, for **(a)** dry (July – September) and **(b)** wet (January – March) periods, between downwind and upwind sites. Gradients were simulated with INV-BRAMS (**red**) and INV-ECMWF (**blue**). Numbers in the y-axis correspond to the observations listed in the respective tables below the plots. Shaded areas on the posterior curves depict the posterior uncertainty for both model estimates. Error bars represent the observation error.

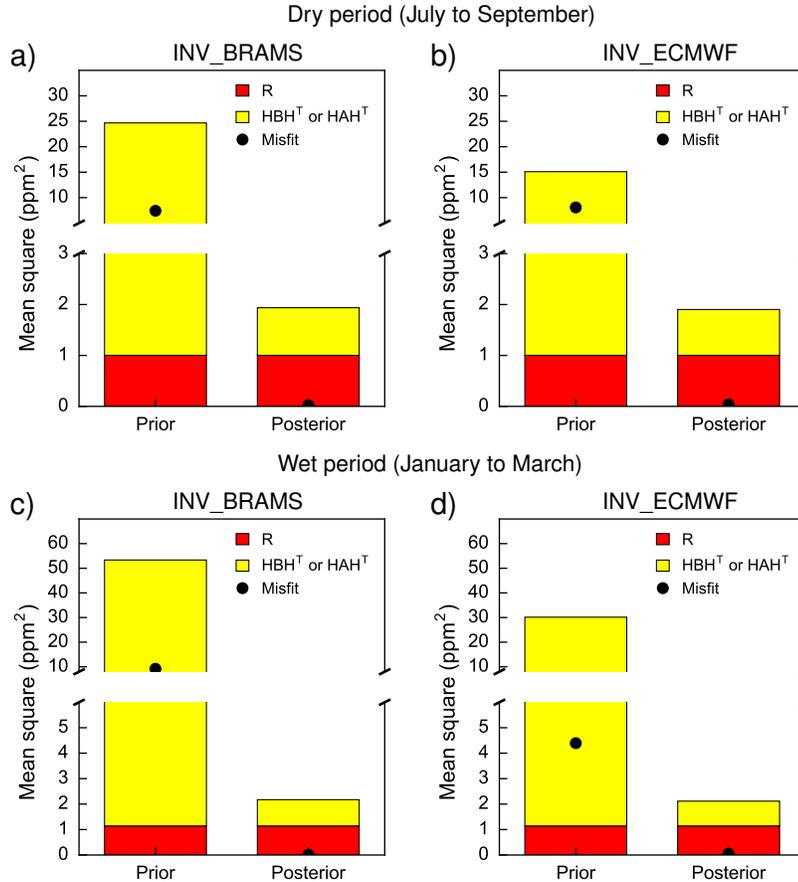


Figure 4.18: Components of Equations 4.6 and 4.7 expressed as mean square (MS) values for the HINTEG inversion. Left-hand terms represented as black dots and right-hand terms as stacked bars.

### 4.3 Results: optimized fluxes

In this section, I present the posterior fluxes for each inversion experiment. At this stage, the previous section showed that different inversions with different observation vectors all had a satisfactory fit to the observations. Here I assess the impact on the fluxes. In this analysis, fluxes are reported such that positive values represent CO<sub>2</sub> emission to the atmosphere (i.e., the surface is a source), and negative values represent removal (i.e., the surface is a sink). In the following subsections, I describe sequentially the results for the four types of inversion experiments detailed in Section 4.1. The results are discussed in Section 4.4. Figures 4.19 and 4.20 report prior and posterior fluxes aggregated over the simulation domain for dry and wet periods, respectively. Figures 4.21 and 4.22 show how the mean NEE was distributed among its component sub-regions. The posterior error correlations are shown in Figure 4.23, where correlations above and below the diagonal correspond to the dry (JAS) and wet (JFM) periods, respectively.

## PROFILE experiment

### Dry season fluxes

Before the inversion, the simulation domain was a net CO<sub>2</sub> source of  $0.63 \pm 1.0$  PgC. After the inversion, the net flux to the atmosphere remained close to the prior in INV-BRAMS ( $0.66 \pm 0.29$  PgC) and increased to  $0.950.30 \pm$  PgC in INV-ECMWF. In both inversions, ocean fluxes remained close to their prior values (Figure 4.19a). This is coherent with the very small sensitivity of concentrations to ocean fluxes (Chapter 3). The significant changes with respect to the prior fluxes can be seen in  $E_{\text{FIRE}}$  and NEE. Prior  $E_{\text{FIRE}}$  of  $0.54 \pm 0.31$  PgC decreased to  $0.24 \pm 0.06$  PgC in INV-BRAMS and to  $0.36 \pm 0.12$  PgC in INV-ECMWF ( $\sim 33 - 55\%$  reduction) (Figure 4.19a). Meanwhile, the NEE increased considerably from a near-neutral prior ( $0.06 \pm 0.98$  PgC) to a net source of  $0.39 \pm 0.29$  PgC with INV-BRAMS, and to  $0.56 \pm 0.31$  PgC with INV-ECMWF. Figure 4.22a shows that at sub-regional level NEE shifted towards a larger source or a smaller sink in all sub-regions, except for regions 2 and 3 in INV-BRAMS, and regions 6 and 7 in INV-ECMWF. All NEE sub-regions showed a large uncertainty reduction of  $\sim 50 - 90\%$  with both models, except for regions 3 and 4, but despite such an uncertainty reduction, the posterior uncertainties remained large ( $> 100\%$  except for sub-region 8), compared to the posterior estimates of the NEE.  $E_{\text{FIRE}}$  sub-regions also showed substantial uncertainty decreased of  $\sim 60 - 86\%$  with the models. Low posterior correlations ( $< 0.1$ ) between  $E_{\text{FIRE}}$  regions E and W (Figure 4.23(a, b)) and moderate posterior uncertainties ( $\sim 30 - 50\%$ ) suggest those two regions were likely well separated. This is possibly due to the location of the sampling sites, where TAB and RBA were likely sensitive to both  $E_{\text{FIRE}}$  regions E and W, while ALF is mostly sensitive to region E. Inversion might have used this information to separate both biomass burning regions. In general, anti-correlations between the posterior uncertainties in different flux sub-regional budgets remained low ( $\leq 0.2$ ), especially between NEE sub-regions. Yet some cases of moderate posterior anti-correlation ( $\sim -0.4 - -0.5$ ) appeared between  $E_{\text{FIRE}}$  and NEE sub-regions with both inversion models (Figure 4.23(a, e)). Therefore, despite the relatively high posterior uncertainties in NEE sub-regions, there were no major flux separation problems between them. The only likely moderate separation problem may be between  $E_{\text{FIRE}}$  sub-region W and NEE sub-region 2.

### Wet season fluxes

Before the inversion, the simulation domain was a net sink of  $-0.47 \pm 1.24$  PgC that turned into a source of  $0.29 \pm 0.31$  PgC with INV-BRAMS and  $0.17 \pm 0.32$  PgC with INV-ECMWF. Similar to the dry period, ocean emissions had a negligible contribution to the budget and were not significantly modified by the inversion (Figure 4.20a). Biomass burning emissions before the inversion were estimated at  $0.20 \pm 0.11$  PgC and remained nearly unchanged after the inversion with both models ( $0.14 \pm 0.09$  PgC for INV-BRAMS;  $0.17 \pm 0.10$  PgC for INV-ECMWF). This is an expected result given the small influence of biomass burning emissions on the concentrations in this period (due to reduced fire activity). The largest changes occurred in the NEE. Before the inversion, the terrestrial biosphere was a strong CO<sub>2</sub> sink of  $\sim -0.70 \pm 1.2$  PgC. After the inversion, however, both models predicted the vegetation was close to neutral ( $\sim 0.10 \pm 0.3$  PgC for INV-BRAMS,  $-0.03 \pm 0.3$  PgC for INV-ECMWF). In both inversion models the largest changes in NEE at sub-regional scale appeared in regions 2 and 8. CO<sub>2</sub> uptake in region 2 increased from  $\sim -0.1 \pm 0.31$  PgC to  $-0.19 \pm 0.12$  PgC and  $-0.39 \pm 0.15$  PgC in INV-BRAMS and INV-

ECMWF, respectively. In region 8, meanwhile, a strong prior sink of  $-0.38 \pm 0.60$  PgC turned into a significant source of  $0.47 \pm 0.18$  PgC and  $0.56 \pm 0.23$  PgC, in INV-BRAMS and INV-ECMWF, respectively. The uncertainty in  $E_{\text{FIRE}}$  sub-regions was not considerably reduced, but this was expected given the reduced fire emissions during the wet period. The uncertainty across the NEE sub-regions decreased by 61 – 92% with INV-BRAMS, and by 51 – 84% with INV-ECMWF, but the posterior uncertainties remained large ( $> 100\%$ ) in some sub-regions. Yet, the low ( $\leq 0.2$ ) posterior anti-correlations suggest that there were no strong separation problems. The only likely moderate separation problem may be in INV-ECMWF, where a moderate posterior anti-correlation ( $\sim -0.5$ ) can be seen between NEE sub-regions 4 and 5 (Figure 4.23 e), both of which exhibited large ( $> 100\%$ ) posterior uncertainties.

## HGRAD experiment

### Dry season fluxes

Prior to the inversion, the simulation domain was a net CO<sub>2</sub> source of  $\sim 0.63 \pm 1.0$  PgC that increased to  $0.87 \pm 0.59$  PgC in INV-BRAMS and to  $0.85 \pm 0.59$  PgC in INV-ECMWF. Thus, posterior fluxes showed a more consistent increase across models than in the PROFILE experiment. Figure 4.20 shows that fire emissions after the inversion remained close to the prior in both models ( $\sim 0.5 \pm 0.3$  PgC). On the other hand, while NEE was close to neutral before the inversion ( $0.06 \pm 0.98$  PgC) it became a CO<sub>2</sub> source of  $0.28 \pm 0.58$  PgC in INV-BRAMS and  $0.36 \pm 0.60$  PgC in INV-ECMWF. There was a fair agreement between the two transport models regarding the direction and magnitude of flux increments across the NEE sub-regions (except for sub-regions 7 and 8, which are both in the south of the domain, less constrained by observations; Figure 4.22). Compared to the PROFILE experiment, the uncertainty reduction in  $E_{\text{FIRE}}$  and NEE sub-regions was smaller when assimilating gradients in both transport models. For  $E_{\text{FIRE}}$  sub-regions the uncertainty reduction was  $\sim 30 - 40\%$  and the posterior uncertainties were moderate to large ( $\sim 45 - 70\%$ ). In NEE sub-regions, the uncertainty reduction varied by  $\sim 6 - 53\%$  in INV-BRAMS, and by  $\sim 3 - 48\%$  in INV-ECMWF, and posterior uncertainties were very large ( $> 100\%$ ) in most sub-regions. In both inversions low ( $\leq 0.2$ ) posterior error anti-correlations appeared between  $E_{\text{FIRE}}$  sub-region W and several NEE sub-regions (1, 2, 5 and 7), and between sub-regions E and 8 (Figure 4.23(b, f)). Therefore even though posterior uncertainty in W remained large ( $\sim 70\%$ ) and posterior uncertainties in NEE sub-regions were also very large, the low anti-correlations suggest the sub-regional budgets of both emission types were rather well separated. An indication of dependence on the transport model is the moderate posterior anti-correlation between sub-regions E and 7 ( $r \sim -0.4$ ) in INV-ECMWF, not seen in INV-BRAMS (Figure 4.23(b, f)). With a moderate posterior uncertainty in sub-region E (45 – 55%) and the large posterior uncertainty in sub-region 7, this may be the only separation problem between these two emission types. The low posterior anti-correlations ( $\leq 0.2$ ) between NEE sub-regions suggest, despite the large posterior uncertainties, an acceptable separation of their respective flux signals.

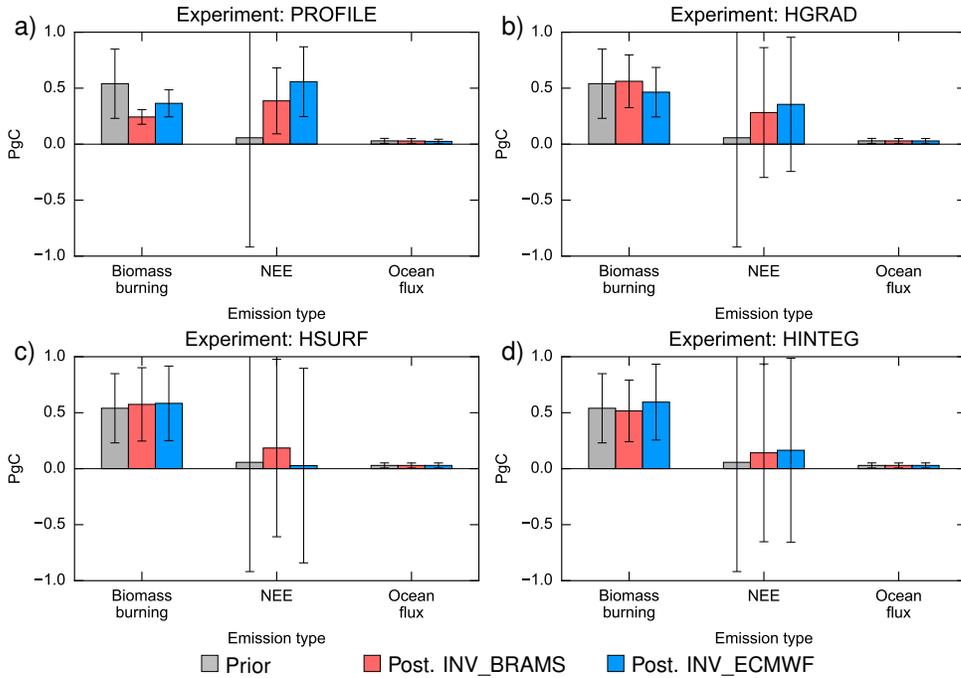


Figure 4.19: Total prior (gray) and posterior biomass burning CO<sub>2</sub> emission, NEE and ocean flux from (red) INV-BRAMS and (blue) INV-ECMWF within the simulation domain for the dry period (July – September). Results from the experiments (a) PROFILE, (b) HGRAD, (c) HSURF and (d) HINTEG. Error bars represent the prior or posterior uncertainty on the fluxes.

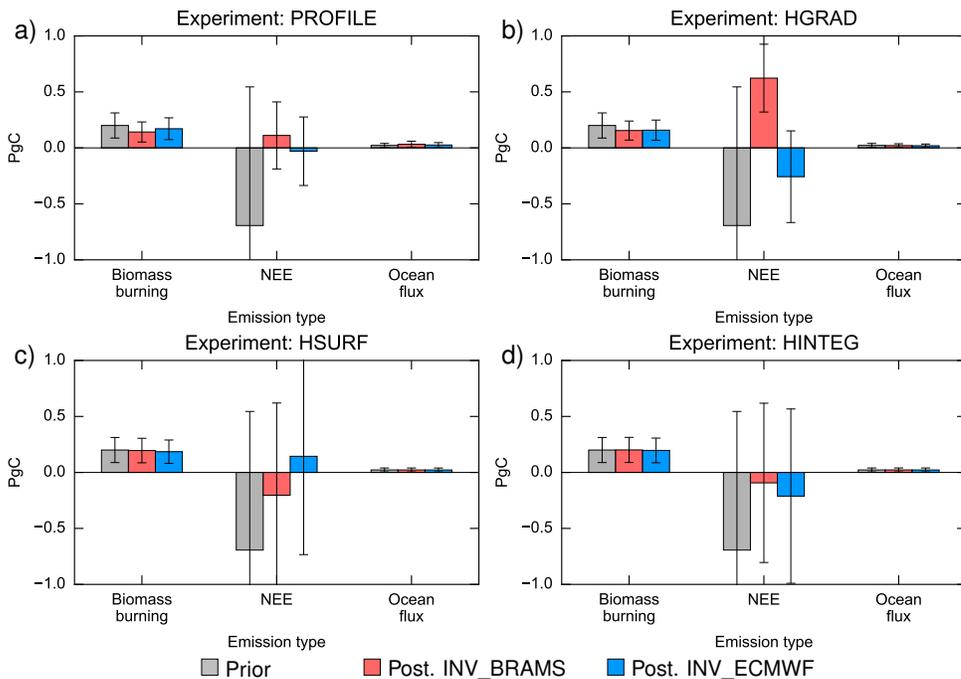


Figure 4.20: Total prior (gray) and posterior biomass burning CO<sub>2</sub> emission, NEE and ocean flux from (red) INV-BRAMS and (blue) INV-ECMWF within the simulation domain for the wet period (January – March). Results from the experiments (a) PROFILE, (b) HGRAD, (c) HSURF and (d) HINTEG. Error bars represent the prior or posterior uncertainty on the fluxes.

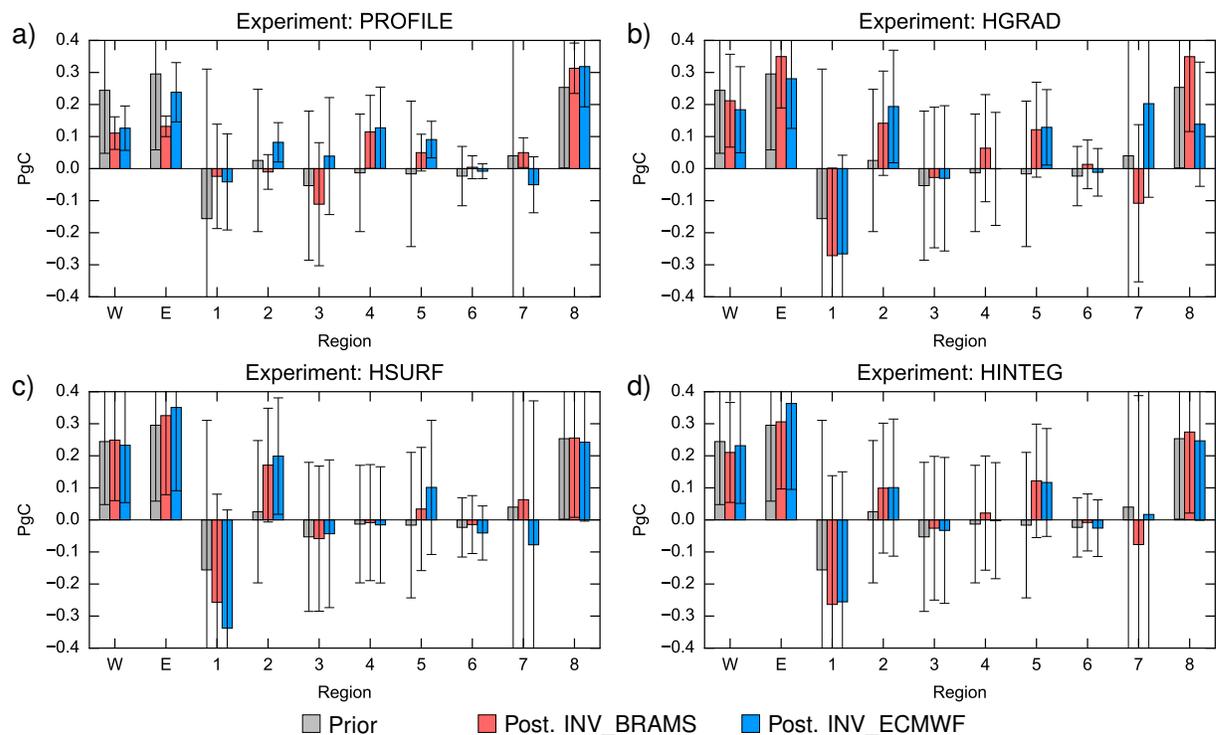


Figure 4.21: Total prior (gray) and posterior sub-regional budget of NEE (1 – 8) and fire emissions (W and E) from INV-BRAMS (red) and INV-ECMWF (blue) for July – September (dry period) 2010. Results for experiments (a) PROFILE, (b) HGRADS, (c) HSURF and (d) HINTEG. Error bars represent prior or posterior uncertainty on the fluxes.

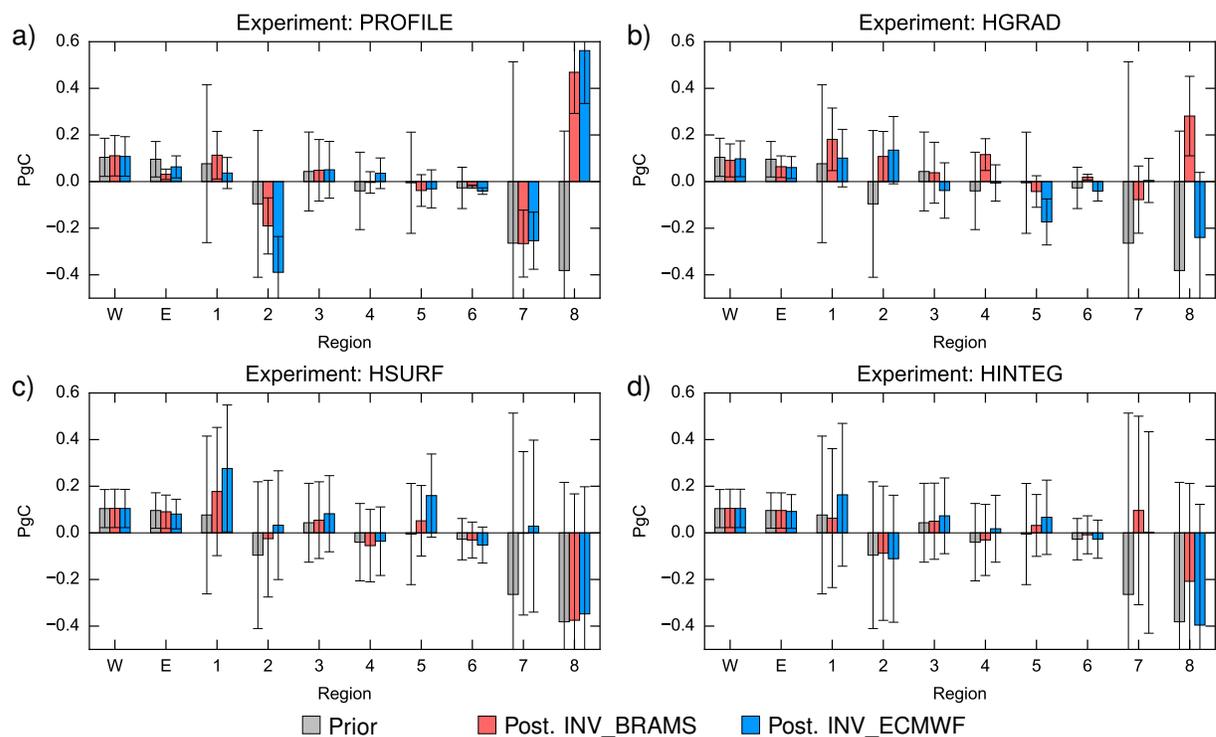


Figure 4.22: Total prior (gray) and posterior sub-regional budget of NEE (1 – 8) and fire emissions (W and E) from INV-BRAMS (red) and INV-ECMWF (blue) for January – March (wet period) 2010. Results for experiments (a) PROFILE, (b) HGRADS, (c) HSURF and (d) HINTEG. Error bars represent prior or posterior uncertainty on the fluxes.

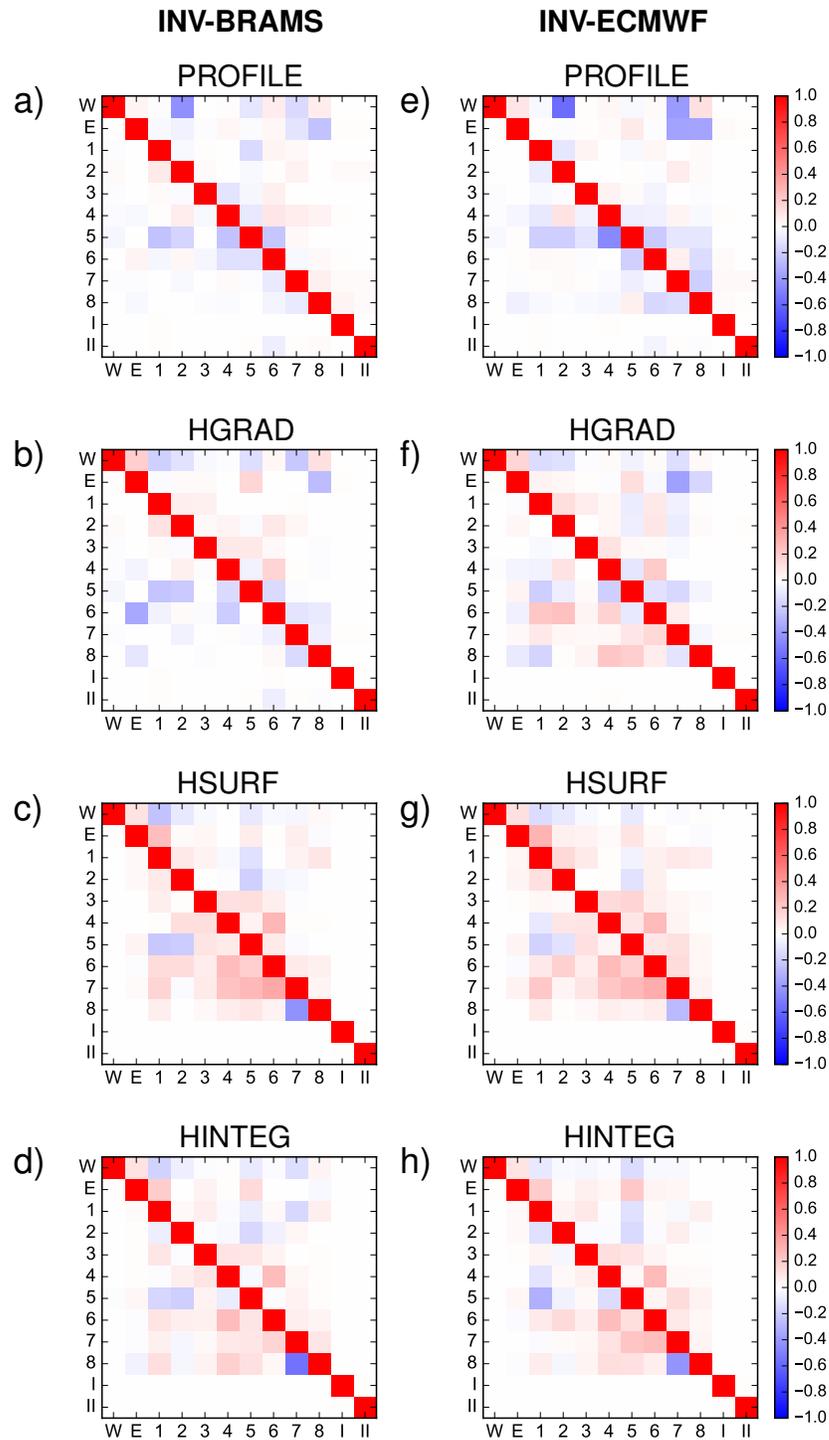


Figure 4.23: Posterior error correlation between NEE regions 1 – 8, biomass burning emission regions W and E, and ocean regions I and II from PROFILE, HGRAD, HSURF and HINTEG inversions, with INV-BRAMS and INV-ECMWF. **Values above the diagonal** represent correlations for July – September (dry period) and **below the diagonal** represent correlations for January – March (wet period).

### Wet season fluxes

Inversions with both transport models diverged considerably regarding the CO<sub>2</sub> budget in the simulation domain. INV-BRAMS turned the strong prior sink of  $-0.47 \pm 1.24$  PgC into a net source of  $0.80 \pm 0.31$  PgC. In INV-ECMWF the domain was almost neutral ( $-0.08 \pm 0.42$  PgC). Posterior  $E_{\text{FIRE}}$  emissions remained very close to the prior value but they were small in this period and should not impact the CO<sub>2</sub> budget. Thus, changes to the domain's budget were caused by a strong uptake reduction by the vegetation (Figure 4.20). Before the inversion, the vegetation was a sink of  $-0.69$  PgC that was reduced to  $-0.26 \pm 0.41$  PgC in INV-ECMWF, and turned into a net source of  $0.62 \pm 0.30$  PgC in INV-BRAMS. This contrasts with the PROFILE experiment, where both models predicted a nearly neutral NEE ( $0.11 \pm 0.3$  PgC in INV-BRAMS,  $-0.03 \pm 0.31$  PgC in INV-ECMWF). Figure 4.22b shows that except in region 5 INV-BRAMS increased the biogenic flux to the atmosphere across all sub-regions, and notably in region 8, in contrast to INV-ECMWF, which increased the uptake in other sub-regions (e.g. regions 3, 5, 6). Both inversion models had a similar performance concerning uncertainty reduction ( $\sim 23 - 86\%$  in INV-BRAMS;  $\sim 30 - 87\%$  in INV-ECMWF), and uncertainty reduction was higher than in the dry period. Low posterior error anti-correlations ( $\leq 0.1$ ) appeared between NEE and  $E_{\text{FIRE}}$  sub-regions (Figure 4.23 b, f); thus, even though posterior uncertainties remained large in both  $E_{\text{FIRE}}$  ( $\sim 70\%$ ) and NEE ( $> 100\%$ ) sub-regions, the low anti-correlations suggest a good separation of NEE and  $E_{\text{FIRE}}$  sub-regional budgets. Low posterior error anti-correlations ( $r \leq 0.2$ ) between NEE sub-regions also appeared, but even though the posterior uncertainties remained large (even after an overall large uncertainty reduction), the anti-correlations were low, such that the flux separation is also acceptable.

## HSURF experiment

### Dry season fluxes

The prior estimate indicated that  $0.63$  PgC were emitted during the dry period in the simulation domain. After the inversion, emissions increased to  $0.79 \pm 0.87$  PgC in INV-BRAMS ( $\sim 26\%$  increase), but only to  $0.64 \pm 0.96$  PgC in INV-ECMWF ( $\sim 2\%$  increase).  $E_{\text{FIRE}}$  emissions increased less than  $10\%$  in both models (Figure 4.19c). The largest change was in the NEE with INV-BRAMS, which increased from  $0.06$  PgC to  $0.18 \pm 0.79$  PgC, while it remained almost neutral in INV-ECMWF ( $0.03 \pm 0.87$  PgC). Figure 4.21c indicates that except for NEE sub-region 7 both models applied similar flux corrections across NEE sub-regions. However, in general the increments were small, coherent with small uncertainty reduction across the NEE regions with both models. Uncertainty decreased by  $\sim 2 - 15\%$  with both inversion models, except in regions 1 and 2, where it was slightly higher ( $\sim 18 - 28\%$ ), and it was more moderate than in the PROFILE experiment. As a result, the inversions yielded large posterior uncertainties in both  $E_{\text{FIRE}}$  ( $\sim 80\%$ ) and NEE sub-regions ( $> 100\%$ ). But as in the HGRAD experiment, despite the large posterior uncertainties, the low posterior error anti-correlations ( $r \leq 0.2$ ) between  $E_{\text{FIRE}}$  and NEE sub-regions (Figure 4.23(c, f)) suggest a fair separation of the  $E_{\text{FIRE}}$  and NEE sub-regional budgets. Similarly, even though posterior uncertainties in NEE sub-regions were very large ( $> 100\%$ ), the low posterior anti-correlations ( $r \leq 0.2$ ) (Figure 4.23(b, c)) between them do not suggest a major separation problem.

### Wet season fluxes

The prior fluxes indicated a strong CO<sub>2</sub> sink of  $-0.47 \pm 1.24$  PgC in the simulated domain. After the inversion the domain was nearly neutral in INV-BRAMS ( $-0.01 \pm 0.83$  PgC), whereas it was a net CO<sub>2</sub> source of  $0.35 \pm 0.89$  PgC in INV-ECMWF. In both models, the uptake reduction is due to changes in the NEE (Figure 4.63). INV-BRAMS reduced vegetation uptake from  $-0.69 \pm 1.24$  PgC to  $-0.20 \pm 0.82$  PgC and INV-ECMWF indicates the terrestrial ecosystem was a net source of  $0.14 \pm 0.88$  PgC. Both models reduced the terrestrial sink in all sub-regions (Figure 4.22c), except in NEE sub-regions 4 and 6, where the sink increased slightly. Uncertainty reduction was, in general, higher than in the dry period across the NEE sub-regions and varied from 2 – 3% (region 3) up to 53 – 55% (region 7), but the posterior uncertainties remained large ( $\sim 80\%$ ) in E<sub>FIRE</sub> sub-regions, and even larger ( $> 100\%$ ) in NEE sub-regions. In spite of these large posterior uncertainties, the nearly null posterior anti-correlations between E<sub>FIRE</sub> and NEE sub-regions, and the low ( $\leq 0.2$ ) anti-correlations between NEE sub-regions themselves (Figure 4.23(c, f)), both suggest that there were no separation problems.

### **HINTEG experiment**

#### Dry season fluxes

The total CO<sub>2</sub> flux to the atmosphere within the simulation domain increased by  $\sim 10\%$  in INV-BRAMS ( $0.69 \pm 0.84$  PgC), and by  $\sim 26\%$  in INV-ECMWF ( $0.79 \pm 0.92$  PgC), relative to the prior value. In INV-BRAMS the changes are attributed to a 4% reduction in E<sub>FIRE</sub> emissions, and to an increase in CO<sub>2</sub> release by the ecosystem ( $0.14 \pm 0.79$  PgC). On the other hand, in INV-ECMWF the changes were due to a 10% increase in E<sub>FIRE</sub>, relative to the prior value, and to an increase in CO<sub>2</sub> release by the ecosystem, which increased about twofold relative to the prior value ( $0.16 \pm 0.82$  PgC) (Figure 4.19d). Figure 4.21d shows that while E<sub>FIRE</sub> and NEE budgets remained about their prior values, significant increments were applied to sub-regional NEE budgets (e.g. sub-regions 1, 2 and 5). The overall low uncertainty reduction (1 – 26%) in the sub-regional budgets resulted in large posterior uncertainties for E<sub>FIRE</sub> ( $\sim 80\%$ ) and NEE ( $> 100\%$ ) sub-regions. Despite such large posterior uncertainties, the low ( $< 0.20$ ) posterior anti-correlations that appeared between E<sub>FIRE</sub> and NEE sub-regions, and between NEE sub-regions (Figure 4.23(d, h)), suggest an acceptable separation of their signals.

#### Wet season fluxes

The domain was a CO<sub>2</sub> sink of  $\sim -0.5$  PgC before the inversion, and became a net source of  $0.13 \pm 0.72$  PgC in INV-BRAMS and  $0.1 \pm 0.79$  PgC in INV-ECMWF. Coherent with the other three inversion experiments, the change is attributed to the terrestrial ecosystem. Before the inversion the vegetation was a sink of  $-0.69 \pm 1.24$  PgC that was reduced to  $-0.09 \pm 0.71$  PgC in INV-BRAMS and to  $-0.21 \pm 0.78$  PgC in INV-ECMWF (Figure 4.20d). The most significant changes to the terrestrial CO<sub>2</sub> sink with respect to the prior were distributed over regions NEE sub-regions 7 and 8 (Figure 4.22d). However, while both models reduced the sink in sub-region 7 such that it is either neutral (INV-ECMWF) or a net source (INV-BRAMS), in region 8 the models disagreed; INV-BRAMS reduced the sink in  $\sim 45\%$  (from  $-0.38 \pm 0.60$  PgC to  $-0.21 \pm 0.42$  PgC), but slightly increased with INV-ECMWF ( $\sim 4\%$ ) (Figure 4.64d). Both inversion models yielded a low to mod-

erate ( $\sim 3 - 48\%$ ) uncertainty reduction across NEE sub-regions, which resulted in large ( $> 100\%$ ) posterior uncertainties. Posterior uncertainties in  $E_{\text{FIRE}}$  sub-regions were also large ( $\sim 80\%$ ). Yet, the nearly null posterior anti-correlations between  $E_{\text{FIRE}}$  and NEE sub-regions (Figure 4.23(e, g)) suggest a good separation of their signals. Likewise, despite the very large posterior uncertainties in NEE sub-regional budgets, the low ( $\leq 0.2$ ) posterior error anti-correlations between them do not suggest major separation problems. The only moderate separation problem appeared due to the moderate ( $0.4 - 0.5$ ) posterior error anti-correlation between NEE sub-regions 7 and 8, where the uncertainties exceed 100

## 4.4 Discussion

I have presented inversion systems using two atmospheric models that assimilated different types of observation to constrain NEE,  $E_{\text{FIRE}}$ , and ocean CO<sub>2</sub> fluxes over tropical South America, during dry and wet seasons. Overall, inversions based on both transport models showed good performance in reproducing the observations under both dry and wet conditions. However, this contrasts with the analysis of Chapter 3, which highlighted potential limitations of the transport models to reproduce the measurements' spatial and temporal variability. Therefore, rather than contradicting the assumptions implied by the diagnostics of Chapter 3, the performance of the inversions may imply that the measurements have been over-fitted, given the transport modeling skills, such that a large portion of the transport model errors influenced the inverted NEE. Diagnostics of the O–B and O–A differences in the observation space indicated that the prior uncertainties, and possibly the observation errors, were overestimated in the set-up of the inversion. If only the prior uncertainties had been overestimated, or if they had been “more” overestimated than the observation errors, this still could have led to data over-fitting. Even though such diagnostics are not fully reliable, since they depend on the theoretical background of the inversion and on the assumptions I made regarding the structure of correlations in  $\mathbf{B}$  and  $\mathbf{R}$ , the overestimation of  $\mathbf{B}$  is highly plausible. In this study, prior fluxes were assigned 100% uncertainty in all flux components (i.e.  $E_{\text{FIRE}}$ , biosphere and ocean fluxes), larger than those in [van der Laan-Luijkx et al. \(2015\)](#), who assigned 80% uncertainty to land fluxes, and 40% to ocean fluxes. As to the observation errors, for the aircraft data [Alden et al. \(2016\)](#) and [van der Laan-Luijkx et al. \(2015\)](#) assigned values of  $\sim 1 - 7$  ppm and 2 ppm, respectively, whereas in this study  $\mathbf{R}$  was built with standard deviations of up to  $\sim 1 - 5$  ppm. In this study  $\mathbf{R}$  was directly derived from the comparison between CH-BRAMS and CH-ECMWF using the same input fluxes, ignoring error correlations, which, in principle, should more easily lead to underestimate rather than overestimate the transport error. An overestimation may come from the high amplitude of the fluxes used as input to the transport models for this comparison. In a general, the consistency diagnostics of [Desroziers et al. \(2005\)](#) suggest the need of tuning the setup of prior and observation errors, but similar to [van der Laan-Luijkx et al. \(2015\)](#), no optimization was performed to assign prior and observation errors. A critical impact of the potential overestimation of  $\mathbf{B}$  and  $\mathbf{R}$  would be a strong overestimation of the posterior uncertainties. Indeed, posterior uncertainties in the inverted fluxes were very large. The large amplitude can be related to the lack of data, even in the PROFILE experiments, where the time lag between profiles of assimilated data at each of the four measurement sites is no less than  $\sim 2$  weeks. But the interpretation of the absolute values of posterior uncertainties requires caution. Thus, in the following I will rather focus on the relative difference between the posterior uncertainties when assimilating horizontal

gradients vs. those in the PROFILE experiments. Such difference was strongly driven by the reduction in the number of data when selecting gradients as a function of the wind direction, rather than assimilating all the individual CO<sub>2</sub> measurements.

For both inversion models, different observation vectors resulted into different CO<sub>2</sub> budgets for the simulation domain (referred to as ‘the region’, henceforth). Likewise, the CO<sub>2</sub> components, i.e. E<sub>FIRE</sub>, NEE and ocean fluxes, were constrained differently depending on the type of observation assimilated. During the dry period, fire emissions remained close to the prior in all experiments assimilating CO<sub>2</sub> gradients, but not in the PROFILE experiment. During the wet period, posterior E<sub>FIRE</sub> emissions were more similar between the different observation vectors, but these emissions were also small and had low impact on the observations. The sub-regional NEE budgets varied considerably between the inversions for both the dry and wet periods. In order to analyze to which extent the differences in NEE and E<sub>FIRE</sub> emissions between the PROFILE and the gradient-based inversions could be attributed to the simple removal of data through the selection of gradients (depending on the wind direction), rather than to a different filtering of the NEE or E<sub>FIRE</sub> signals from the data by the inversion when assimilating gradients, flux corrections to E<sub>FIRE</sub> across all sub-regions and weeks during the dry season were compared for PROFILE and HGRAD experiments for INV-BRAMS (Figure 4.24).

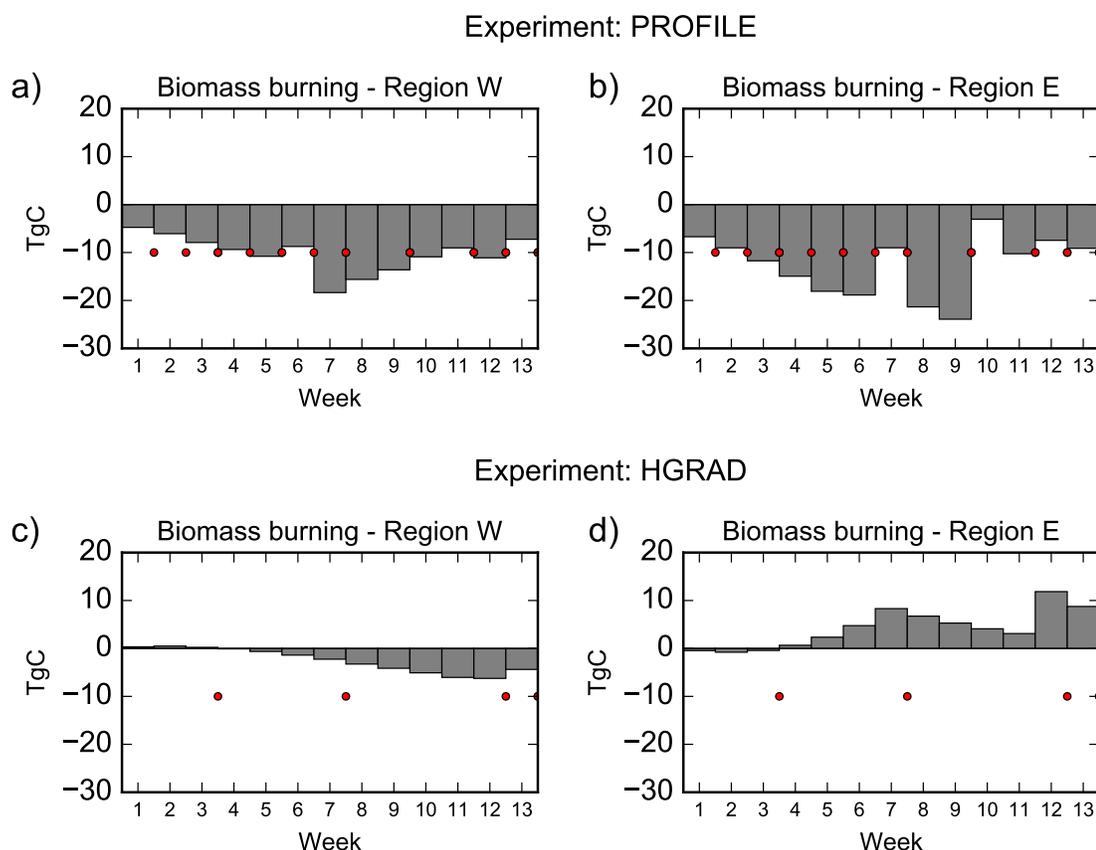


Figure 4.24: Impact of the observations assimilated in the (a, b) PROFILE and (c, d) HGRAD experiments on the posterior weekly, sub-regional biomass burning CO<sub>2</sub> emissions, for the July – September (dry period) for INV-BRAMS. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observation were available.

#### 4. REGIONAL ATMOSPHERIC INVERSION OF CO<sub>2</sub> IN AMAZONIA

For  $E_{\text{FIRE}}$  on sub-region W, flux increments in PROFILE (Figure 4.24a) and HGRAD (Figure 4.24c) had different magnitude, although they had the same direction, whenever observations were available for the same week. But for sub-region E, when observations were available in the same week, flux increments had opposed direction over most weeks. Such differences can also be seen across NEE sub-regions, for instance, in regions 1 – 3, 7 and 8 (Figures 4.25 and Figure 4.26). These findings were also valid for INV-ECMWF (Figures B.1 to B.3 in Appendix B), despite some differences indicative of transport model dependency. These results support that inversions exploited the information of individual measurements and horizontal gradients differently.

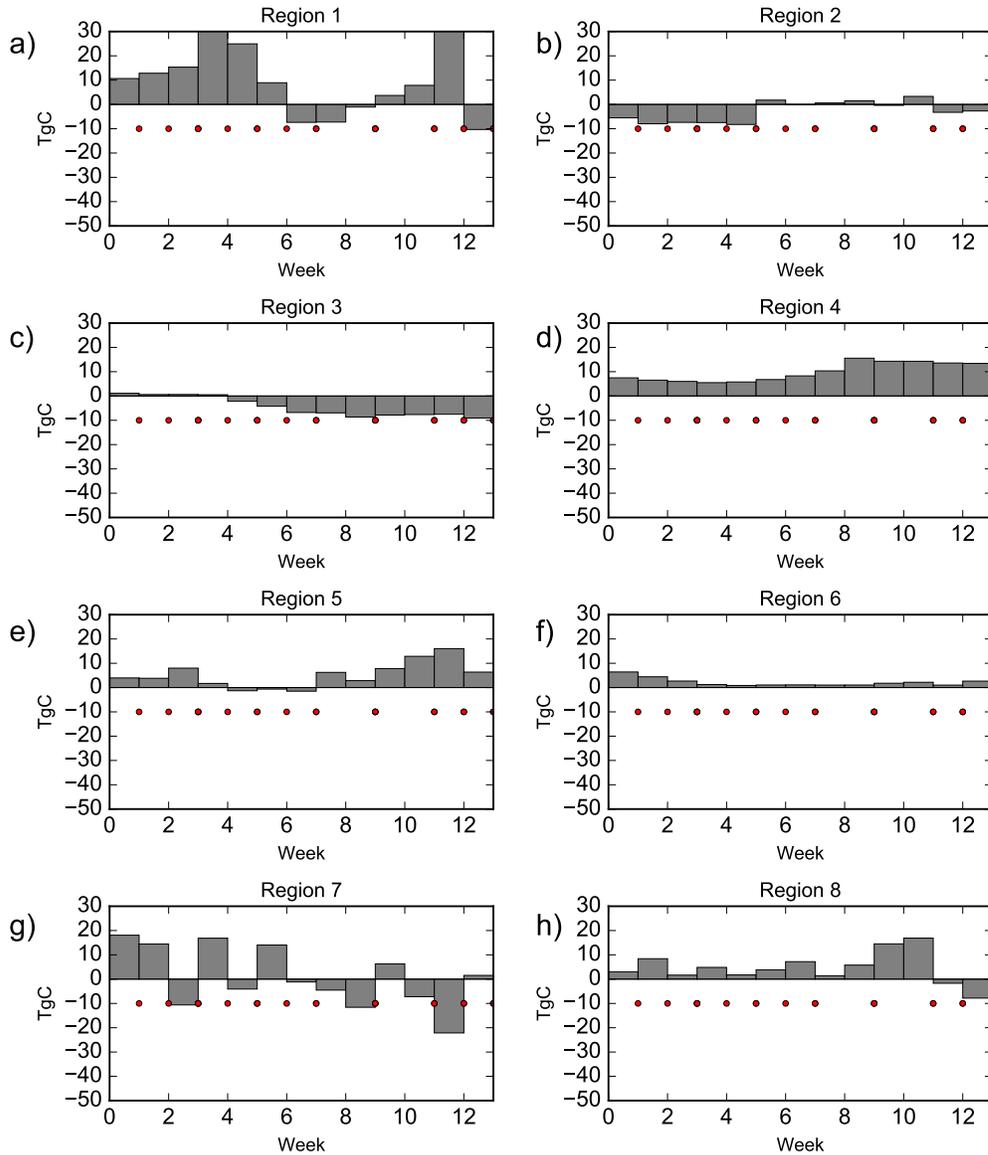


Figure 4.25: Impact of the observations assimilated in the PROFILE experiment on the posterior weekly, sub-regional NEE, for the July – September (dry period) for INV-BRAMS. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observation were available.

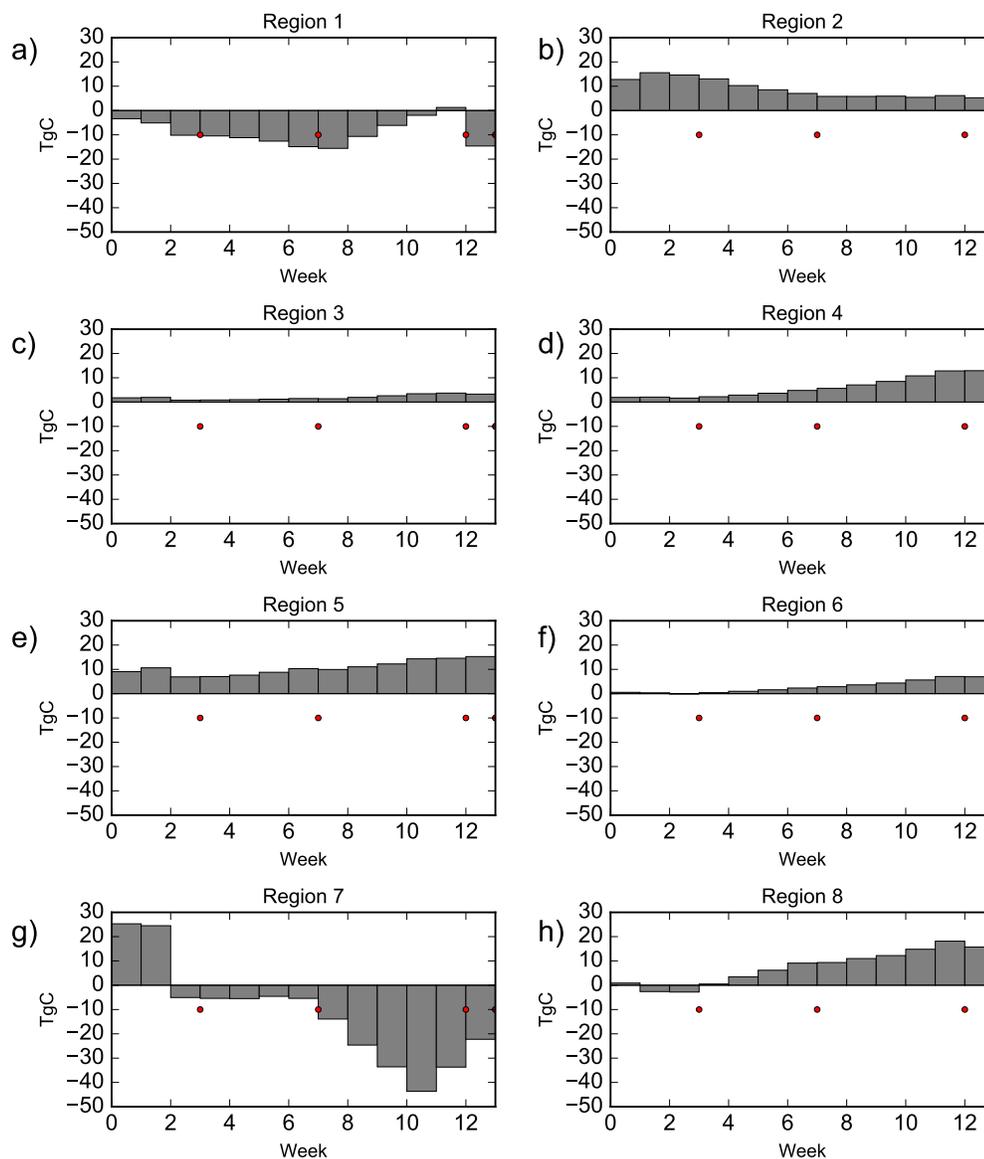


Figure 4.26: Impact of the observations assimilated in the HGRAD experiment on the posterior weekly, sub-regional NEE, for the July – September (dry period) for INV-BRAMS. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observation were available.

Furthermore, during the dry period there was some agreement, both in magnitude and direction, in the posterior sub-regional NEE budgets from gradient-based inversions, but not during the wet period. Differences between gradient-based inversions during the wet period are indicative of the impact of the choice of the observation vector. These analyses demonstrate the need for selecting the most adapted observation vector for the inversion. A poor selection could lead to strong uncertainties in the inverted fluxes that are not well diagnosed by the inversion system. The rather simple models used to build the matrices  $\mathbf{B}$  and  $\mathbf{R}$  in traditional inversion systems hamper the ability to filter the content of the model-observation misfits related to errors in the NEE estimates (which is ignored by

the diagnostic of posterior uncertainties in the inversion system). Selecting an adapted observation vector helps the inversion system filter out other error sources errors and better filter the signature associated with the control vector in the model-observation misfits. However, in my inversions there was a trade-off between the need for increasing the relative signal of the NEE in the observation vector and retaining sufficient data to strongly decrease the posterior uncertainty in the flux estimates—the amount of data being dramatically reduced by the selection of gradients along the wind direction. PROFILE exhibited the largest uncertainty reduction of all inversion experiments during both dry and wet periods, likely because PROFILE had more observations than the gradient-based inversions. The much larger posterior uncertainties obtained when assimilating gradients than in the PROFILE experiment do not give more confidence in the optimized fluxes from the HGRAD, HSURF or HINTEG experiments than in those from the PROFILE inversions. Yet PROFILE is strongly subject to the uncertainties in fluxes outside South America. Gradient-based inversions, on the other hand, are less dependent on the flux budget of regions upwind of the observation sites in South America, not constrained by the data of GA2014. Chapter 3 showed that the signature of CO<sub>2</sub> boundary conditions was virtually suppressed in gradient-based observations. However, horizontal gradients should be calculated between the Amazon sites and ocean sites (like GA2014) in order to constrain fluxes over land comprised between sites ALF and SAN and the coast.

Inversions showed multiple error anti-correlations between NEE sub-regions, or between these and E<sub>FIRE</sub> sub-regions. However, the low anti-correlation values, for instance, in PROFILE experiments, were an indicator of good separation of sub-regional flux budgets. In gradient-based inversions, overall, even though posterior uncertainties were larger than PROFILE, the low error anti-correlations still suggest the sub-regional flux budgets were fairly separated. Only a few cases of potential separation problems appeared between NEE sub-regions, or between these and E<sub>FIRE</sub> sub-regions, due to moderate posterior error anti-correlations and large posterior uncertainties, especially in NEE sub-regions.

Inversions with BRAMS and ECMWF illustrate the uncertainty in the flux estimates, both at regional or sub-regional scale, stemming from the underlying atmospheric model. As suggested in previous inversion studies (Peters et al., 2007; Peylin et al., 2002), the use of multiple inversion estimates provides an additional measure of the uncertainty in the estimated fluxes, just as important as the formal uncertainties derived from the Bayesian approach employed here. The differences between INV-BRAMS and INV-ECMWF support the assumption that the data were over-fitted by the inversion due to the overestimation of the prior uncertainties, and that a large portion of the transport errors were projected into the inverted fluxes.

I could not strictly confront my estimates of the Amazon carbon balance to that of GA2014. My choice of assimilating horizontal gradients between inland sites allowed the inversion system to constrain mainly the fluxes over areas between pairs of measurement sites, leaving the areas between the coast and the sites closest to the coast mostly unconstrained. GA2014, on the other hand, established gradients between ocean stations and the inland sites, which allowed them to strongly constrain fluxes across most of Amazonia. The experiment PROFILE was likely to constrain a large portion of the Amazon basin. Yet GA2014 estimated their carbon balance based on an area calculated from mean annual footprints of the individual measurements at each site, an area which likely differs from my definition of the Amazon basin (Figure 4.1). But in spite of those methodological differences between my study and that of GA2014, I used their results to bring insight on the realism of my flux estimates. I included the results of van der Laan-Luijkx et al. (2015) and

Alden et al. (2016) in the comparison since they also used an inverse modeling approach with the data of GA2014.

I compared my inversion estimates of the NEE within the Amazon basin (as defined in Figure 4.1) and the estimates of the Amazon net biome exchange (i.e. the surface-atmosphere carbon flux excluding fire emissions) from GA2014, van der Laan-Luijkx et al. (2015) and Alden et al. (2016) for January – March and July – September 2010. For the comparison, I averaged the budgets of my gradient-based inversions. In the dry period, my inversions predicted the basin was a net source of CO<sub>2</sub> between  $0.13 \pm 0.49$  and  $0.36 \pm 0.25$  PgC, opposite to the three published studies, which predicted a sink of  $-0.02$  to  $-0.27$  PgC. Despite the spread in my gradient-based inversions, on average they all corrected the prior sink in the same direction. My inversions lay outside the range of the three published estimates, but were consistent in turning the prior CO<sub>2</sub> sink into a net source (Figure 4.27a).

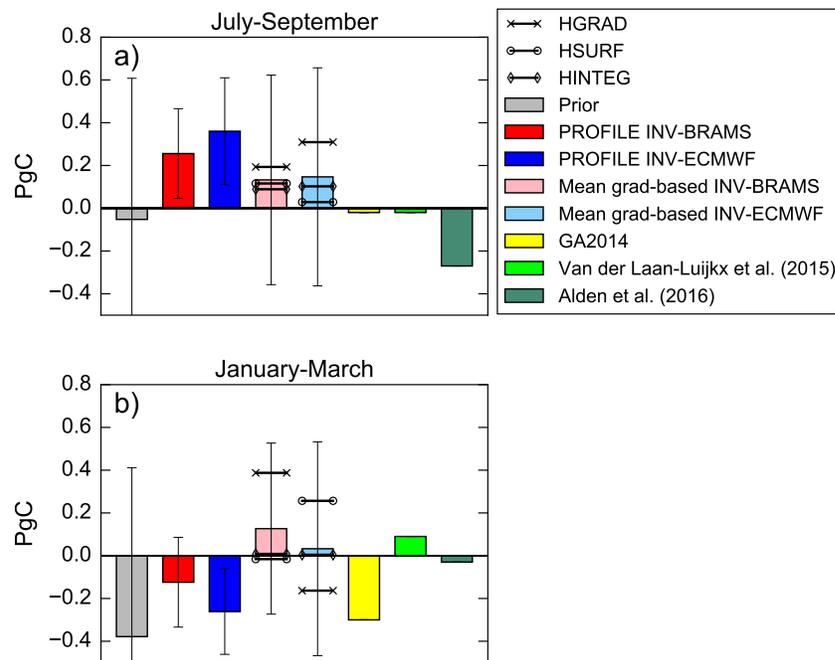


Figure 4.27: NEE budget for the Amazon basin for (a) dry and (b) wet periods. **Grey bar** is the prior NEE budget. **Red** and **Blue** bars are the NEE budget from PROFILE inversions with INV-BRAMS and INV-ECMWF, respectively. **Light red** and **light blue** are the average of the three gradient-based inversions for INV-BRAMS and INV-ECMWF, respectively, and horizontal lines denote the NEE balance for the individual gradient-based inversions. **Yellow bar** is the estimate of GA2014; **light green** and **dark green** are the estimates of van der Laan-Luijkx et al. (2015) and Alden et al. (2016), respectively

During the wet period (Figure 4.27b), all my inversions were coherent with the reduction of the strong prior CO<sub>2</sub> sink. PROFILE inversions predicted the basin was still a sink of  $-0.12 \pm 0.21$  to  $-0.26 \pm 0.20$  PgC, whereas the average of gradient-based inversions predicted a net source  $0.03 \pm 0.50$  to  $0.13 \pm 0.40$  PgC. With both INV-BRAMS and INV-ECMWF, two out of the three gradient-based inversions were consistent with the PROFILE inversions in predicting the vegetation was either neutral or a net sink of CO<sub>2</sub>, like

GA2014 ( $\sim -0.30$  PgC) and Alden et al. (2016) ( $-0.03$  PgC) (Figure 4.27b). It is difficult to conclude about the realism of my flux estimates because, compared to van der Laan-Luijkx et al. (2015) and Alden et al. (2016), my results produced a wide range of NEE estimates, both as a function of the observation vector and of the atmospheric model chosen. Nevertheless, the spread of my results reflects the wide range of results from other studies and my inversions produced a general pattern during the dry season 2010, i.e. the Amazon basin was a strong CO<sub>2</sub> source, which adds to the existing range of estimates from inversions studies in Amazonia.

## 4.5 Conclusions

The application of regional atmospheric inversion in this chapter hardly strengthened the knowledge on the seasonal or inter-annual variability of the NEE in the Amazon basin, even though this should be the long-term goal of such activity. Like in Chapter 3, this chapter rather highlighted the strong need for studies and development to improve the skill of the regional inversion in Amazonia before reaching a level of confidence that is sufficient for the in-depth analysis of the temporal variations of NEE in this region. This study has participated in such a long term effort, and it still contributed to enrich the set of estimates of the impact of the extreme dry season in 2010 on the NEE.

My results were strongly dependent on the atmospheric transport models, and therefore on the transport model errors, even though I used regional mesoscale transport configurations and a dedicated regional meteorological model, or a relatively high-resolution operational forecast assimilating large amounts of data in the area. Chapters 3 and 4 demonstrate the critical need for refining the regional transport modeling to achieve robust regional inversions. This may be one of the main explanations for the large differences in the estimates obtained by previous studies (GA2014; van der Laan-Luijkx et al., 2015; Alden et al., 2016) who based their inversion computations on different transport models or proxies.

The dependency of my results of the inverted fluxes on the choice of the observation vector reveals the associated uncertainty, and could also explain the large differences in the estimates obtained by previous inversion studies in Amazonia (GA2014; van der Laan-Luijkx et al., 2015; Alden et al., 2016) based on different types of observation vectors. Especially during the wet period, there was little agreement in the regional NEE budget among the inversions with different observation vectors. In addition, during the dry period, gradients and individual measurements in profiles constrained the regional budget of E<sub>FIRE</sub> emissions differently. While E<sub>FIRE</sub> emissions remained close to the prior value when assimilating gradients, important corrections were applied when assimilating the individual measurements. Further, even though the number of individual measurements in the profiles was 10 – 20 times higher than the number of gradients available for the inversion, the distribution of flux corrections to NEE and E<sub>FIRE</sub> emissions across sub-regions/weeks in inversions using gradients suggested that the few gradients assimilated were sensitive to most sub-regions/weeks, and that flux corrections were spread to other sub-regions/weeks through correlations in the prior covariance matrix. Therefore, there is some confidence that differences between the flux budgets, when using different observation vectors, could be attributed to the inversions exploiting the observations in different ways, and that it was not an artifact generated by a reduced number of observations.

During the dry season, two elements support that inversions with gradients separated

better the signal of  $E_{\text{FIRE}}$  emissions from that of NEE. First, at least with INV-BRAMS, inversions using gradients exhibited a smaller number of posterior anti-correlations between sub-regions, for both NEE and  $E_{\text{FIRE}}$  emissions, than when using individual measurements. Second, flux corrections and uncertainty reduction in posterior  $E_{\text{FIRE}}$  emissions were lower with gradients than with individual measurements. Therefore, flux corrections to NEE were likely less affected by the uncertainty in  $E_{\text{FIRE}}$  emissions when assimilating gradients. This adds to the advantage brought by horizontal gradients over individual measurements; in Chapter 3 the former showed less sensitivity to the uncertainty in boundary conditions than individual measurements did.

It is difficult to judge the realism of posterior NEE inferred using gradients, given the large posterior uncertainty, even assuming that it could have been overestimated. Therefore, I compared my results with the estimate of GA2014, and with the inverse modeling studies of [van der Laan-Luijkx et al. \(2015\)](#) and [Alden et al. \(2016\)](#). Using GA2014 as a reference, my inversions with both profiles and gradients predicted the vegetation was too high a source of  $\text{CO}_2$  during the dry period. During the wet period, inversions with individual measurements agreed the vegetation was a sink of  $\text{CO}_2$ , and were closest to the estimate of GA2014. Similar differences can be seen between GA2014 and estimates of [van der Laan-Luijkx et al. \(2015\)](#) and [Alden et al. \(2016\)](#), both of which missed the strong sink that GA2014 estimated during the wet period. Further, [Alden et al. \(2016\)](#) agreed with the near-neutral vegetation found by GA2014, while [van der Laan-Luijkx et al. \(2015\)](#) found a large  $\text{CO}_2$  sink. Therefore, the differences between my inversions and GA2014 are in the range of differences with respect to the recent regional inversion studies in Amazonia. The critical element illustrated through my inversions is the large range of NEE estimates that arises from the choice of the transport model and from the selection of the observation vector.



---

## Conclusions and perspectives of future research

The goal of my thesis consisted in improving the knowledge of the seasonal and inter-annual variations of NEE in Amazonia using a regional inversion modeling approach. When I started my thesis, the estimates of the CO<sub>2</sub> balance in Amazonia from atmospheric inversion models relied on global systems. The lack of data from local stations in Amazonia to constrain the flux estimates in the region and global transport modeling shortcomings gave a low confidence on such estimates.

In the first phase of my thesis I looked deeper into the flux estimates from such global inversions than usually done in inter-comparison studies of global inversions. To this purpose, I carried out two global inversions that assimilated data from the global observation network, and new data from four local ground stations were assimilated in one of them; some of those local data had not been used in global inversions. I examined the robustness of seasonal and inter-annual variations of the net ecosystem exchange (NEE; excluding fire emissions) in Amazonia inferred from those two inversions (Chapter 2). The global inversions exposed limitations of the local observation network that existed by the time I started my thesis, as well as the limitations of the global approach. This supported the objective of developing a regional modeling approach. I developed two CO<sub>2</sub> regional transport model configurations which were evaluated with the new airborne CO<sub>2</sub> observations from Gatti et al. (2014). Further, I used the two transport configurations to explore how to exploit airborne measurements, combined into different types of observation vectors of individual measurements or of gradients between bins of these data in the wind direction, to obtain robust estimates of NEE in Amazonia (Chapter 3). There were difficulties to demonstrate that the relative benefit of the observation vectors based on along-wind gradients to separate the uncertainty in NEE from other sources (e.g. biomass burning CO<sub>2</sub> emissions ( $E_{\text{FIRE}}$ ), transport model error) was worth decreasing the number of data to be assimilated, through the definition and selection of these gradients. It justified carrying out inversions with both transport model configurations and with the different types of observation vectors, to assess the robustness and consistency of the seasonal and year-to-year variations of NEE inferred from the inversion (Chapter 4).

In this context, I present the main conclusions of my thesis works (Section 5.1), and perspectives of where my research could be extended (Section 5.2).

## 5.1 Conclusions

### **The ability of a coarse resolution inversion to estimate seasonal and inter-annual variations of NEE in Amazonia**

The first phase of this thesis showed that in a coarse resolution global inversion the information from local ground-based stations, as well as that from stations distant from Amazonia, controlled the NEE in tropical South America at a scale not adapted for solving the seasonal and inter-annual variations of NEE in Amazonia. This control was characterized by alternate (dipole) areas of positive and negative flux corrections, both over land and over the Southern Hemisphere Ocean. The local stations data mainly modified the amplitude of the dipole over land and shifted its position but did not reduce it, even though they generated some corrections in the vicinity of the stations.

Even though the inversions using the new ground-based sites improved the seasonal variations of the simulated CO<sub>2</sub> concentrations at these sites, the resulting seasonality of NEE in tropical South America was not substantially different from that of the inversions ignoring these sites. Furthermore, over rain forests, which cover most of tropical South America, the seasonality of inverted NEE was not in agreement with the assumption of stronger CO<sub>2</sub> uptake during periods of higher insolation suggested from observation products derived from flux towers. Despite some more robust results like the indication of a net source ( $\sim 0.21$  PgC) during 2005, year that was marked by a severe drought, and the strong sink in 2009 ( $\sim 1.1$  PgC), when abnormal humid conditions prevailed, these elements reduced the reliability on the seasonal and inter-annual variability of NEE inferred through the inversions, and supported that the patterns of flux corrections across the Amazon basin mostly reflected larger scale corrections from the global inversion across the southern hemisphere, with artificial impacts regionally. This was a consequence of the sparsity of the local network and the limited measurement records at the local sites, but also as a consequence of the coarse resolution of the inverse modeling framework, which was inherent to its global configuration.

### **Regional atmospheric modeling of CO<sub>2</sub> transport in Amazonia**

I devoted the second phase of my thesis to assessing the potential benefit of the regional atmospheric model BRAMS for CO<sub>2</sub> transport simulations in Amazonia, relative to using the global system ECMWF, when both models were used to force the tracer-transport model CHIMERE. To this purpose, I developed a new offline interface between BRAMS and CHIMERE, which provided a new regional meteorological-transport configuration for Amazonia. BRAMS has been tailored to improve key transport processes in Amazonia, it can generate meteorology at high spatial resolution, and it can be calibrated for specific studies in the region. ECMWF is, on the other hand, an operational model that assimilates meteorological data from diverse data streams. As a first indicator of the quality of the simulated transport fields, the validation of the meteorology generated with both models showed a similar performance in terms of the statistics of the model-observation misfits, and the performance of both models was similar to that seen in previous studies of validation of meteorological models in the region.

CO<sub>2</sub> transport simulations with both model configurations were compared to vertical profiles of CO<sub>2</sub> observations at four sites in Amazonia. Both meteorology-transport configurations showed, again, similar performance. Instead of demonstrating the superiority of the regional, transport-meteorological configuration CHIMERE-BRAMS, the com-

parisons between CHIMERE-BRAMS and CHIMERE-ECMWF highlighted and characterized the high impact of transport modeling uncertainties, even when using these two relatively high-resolution model configurations for modeling vertical profiles of CO<sub>2</sub>. These elements justified using both configurations for regional inversions in the last phase of my thesis.

The sensitivity of the simulated vertical profiles to the meteorological forcing and to the different flux components did not indicate evidently a “best” transport configuration. Yet, the sensitivity tests proved useful to explore the ability of an inversion system based on these two transport configurations to separate the signature of the uncertainty in NEE from that in the other transport components when assimilating four different types of observation: the individual measurements, the horizontal along-wind gradients between pairs of measurement profiles for five vertical layers, the horizontal along-wind gradients in the surface layer between pairs of profiles sites, or horizontal along-wind gradients between the vertically integrated profiles. The four types of observation were highly sensitive to NEE and uncertainties in this flux, but also to  $E_{\text{FIRE}}$  emissions in the dry season and to transport uncertainties. They were much less sensitive to boundary conditions, while the influence of the fossil fuel emissions and ocean fluxes was almost negligible. Thus, inversions using aircraft data can potentially improve our knowledge of the NEE, without strong influence from the uncertainty in the other flux components, with the notable exception of biomass burning emissions during the dry season, and with a strong requirement of having accurate regional transport modeling.

The analysis of horizontal along-wind gradients of CO<sub>2</sub> between profiles to evaluate the two transport configurations exhibited two advantages compared to the evaluation against individual profile measurements in view to infer the NEE through inversion. Using the gradients to evaluate the models nearly suppressed the influence of the boundary conditions, and showed lower transport model differences either close to the surface or in altitude. Moreover, in horizontal along-wind gradients, while the impact of uncertainties in both transport model and NEE decreased with altitude, the uncertainty in the transport model became less significant than that in NEE in altitude, suggesting that the transport error does not mask the uncertainty in NEE in altitude as much as it does close to the surface. Therefore, assimilating horizontal along-wind gradients, particularly considering different vertical layers, seemed to offer a stronger potential than individual measurements to constrain the NEE with a limited influence of the uncertainties in boundary conditions, and, to a lesser extent, from the transport model uncertainty. The uncertainty in  $E_{\text{FIRE}}$  emissions, though, had a similar weight, relative to that in NEE, in the four observation vectors. On that ground, it was unclear that horizontal along-wind gradients would make inversion estimates more robust than individual measurements, while their definition implies a strong selection and binning of the data to be assimilated, and thus, potentially, a loss of useful information to constraint NEE in the inversion.

### **Regional inversion of CO<sub>2</sub> fluxes in Amazonia**

In the third phase of my thesis, I developed two analytical, regional inversion systems for tropical South America to infer the regional and sub-regional flux budgets of NEE,  $E_{\text{FIRE}}$  emissions and ocean fluxes. The two inversion systems were based on the two meteorology-transport model configurations—CHIMERE-BRAMS and CHIMERE-ECMWF—set up and evaluated in Chapter 3. With the two systems and with the same set of aircraft CO<sub>2</sub> measurements as in Chapter 3, I made inversions with the four types of observation vectors introduced in Chapter 3, i.e. the individual measurements at each

profile site and horizontal along-wind gradients between pairs of sampling sites at different altitudes, close to surface or the vertical integral.

The inverted regional and sub-regional flux budgets of NEE or  $E_{\text{FIRE}}$  emissions were strongly dependent on the choice of the transport model, illustrating the importance of the transport modeling errors. Transport model uncertainty is likely one of the main reasons for the large differences between the published regional inversion estimates of the Amazon carbon balance, which have relied on different models or approaches to simulate the transport in the region. My study highlights that improving transport models is essential to yield robust estimates.

Both regional and sub-regional budgets of NEE and  $E_{\text{FIRE}}$  emissions were also strongly dependent on the selection of the observation vector. This was evident particularly for NEE during the wet period and for  $E_{\text{FIRE}}$  emissions in the dry period. During the wet period, there was less agreement on the regional budget of NEE from the inversions with the four observation vectors than during the dry period. In the dry period, while inversions with gradients did not modify substantially the budget of  $E_{\text{FIRE}}$  emissions, the inversion assimilating individual measurements introduced significant corrections. The distribution of flux corrections across sub-regions/weeks illustrated that when individual measurements and gradients were available for the same week, the inversion introduced different flux corrections. These elements supported that the differences between the flux budgets inferred with individual measurements or horizontal gradients were not a consequence of the number of observations available in the observation vector, but a result of the way the inversions exploited the different types of observation.

Chapter 3 illustrated that gradient-based observations were less sensitive to uncertain fluxes outside the simulation domain than inversions with individual measurements. On top of that, in Chapter 4 gradient-based inversions showed a better separation of the flux budgets of NEE and  $E_{\text{FIRE}}$  emissions in the dry period than inversions with individual measurements. The former yielded fewer anti-correlations between the posterior uncertainties in NEE and  $E_{\text{FIRE}}$  sub-regions, and lower flux corrections and posterior uncertainty reduction in the regional budget of  $E_{\text{FIRE}}$  emissions than the latter. As a result, flux corrections to NEE were likely less influenced by the uncertainty in  $E_{\text{FIRE}}$  emissions in gradient-based inversions.

The different inversions diagnosed a large posterior uncertainty in NEE, and thus a low confidence for tropical South America in their optimal estimate of the NEE. However, I compared this posterior NEE from my inversions to the  $\text{CO}_2$  flux estimate (excluding fire emissions) for the Amazon basin from [Gatti et al. \(2014\)](#). My inversions based on gradients or individual profile measurements predicted a higher source of  $\text{CO}_2$  during the July – September (dry period) than [Gatti et al. \(2014\)](#). During January – March (wet period), the inversions with individual measurements were closest to the strong sink estimated by [Gatti et al. \(2014\)](#). In spite of these discrepancies, my results were in the range of differences between [Gatti et al. \(2014\)](#) and the recent inverse modeling studies of [van der Laan-Luijkx et al. \(2015\)](#) and [Alden et al. \(2016\)](#). In view of this comparison, as well as their own diagnostics of uncertainties in the inverted NEE, my regional inversion estimates hardly bring new information on the seasonal and year-to-year variations of NEE in Amazonia. Yet, my study contributes to the long-term effort to improve our knowledge of the Amazon carbon balance. The study highlights the need for improving not only transport modeling and the observation network in the region, but also the inverse modeling strategy, through, at least, a better definition of the observation vector to account for the specificities of the available measurements and the weaknesses of current transport models in Amazonia.

## 5.2 Perspectives

This thesis, and the regional inversions that were developed in parallel by other research groups, illustrate that large uncertainties remain in the estimates of seasonal and year-to-year variations of the carbon balance in Amazonia. It is necessary to continue to improve the regional CO<sub>2</sub> transport modeling skills. In particular, the simulation of the regional meteorology, either with BRAMS or with other regional atmospheric models should be improved with the aim of achieving better performances than with global systems like ECMWF. Likewise, despite the recent efforts to sample the atmosphere within and around Amazonia, the definition of the observation vector, and the underlying selection of the data to be assimilated, have a strong influence on the inferred fluxes. Thus, this research should be oriented to strengthen the strategy for assimilating data, and in particular aircraft data, which still represent the most important data source in Amazonia.

The recurrent limitation of atmospheric inversions is the amount of data to constrain the fluxes, and airborne observations every two weeks (e.g. [Gatti et al., 2014](#)) may not be enough to reduce the uncertainty in the inferred fluxes. While traditional atmospheric inversions usually avoid assimilating aircraft profiles, the current network configuration in Amazonia makes it necessary. Moreover, these data, in principle, should present a strong advantage compared to ground-based in situ measurements since aircraft profiles provide an essential view of the vertical structure of the CO<sub>2</sub> distribution in a region where the deep convection plays a critical role in the CO<sub>2</sub> variability. Chapters 3 and 4 have shown that these profiles bear patterns, even at heights higher than 2 km above ground level, whose temporal representativeness is likely very short, which can hardly be fitted and interpreted with the atmospheric transport models I used and thus strongly influence the inversion fluxes. In continuous ground-based measurements, such patterns can be characterized and filtered, or at least their impact can be limited, by giving more weight to the information at large temporal scales in the inversion frameworks. This is far more difficult when analyzing vertical profiles every two weeks. Increasing the frequency of aircraft data, though, may be limited by economic reasons. More studies will be needed to refine the strategy for assimilating aircraft profiles to take better advantage of their vertical information, and account for the limitation of their relatively low frequency. Improving the modeling of the vertical transport will be required for this; which mainly means analyses will have to be oriented in the direction of what has been investigated in Chapters 3 and 4. Furthermore, it would also be important to assimilate complementary data. In situ ground measurements have been scarce in the region, but efforts have been made to install new equipment. Continuous analyzers to measure CO<sub>2</sub> near the surface are currently installed at Santarém (Brazil) and French Guyana, but these measurements are just above the forest canopy and may not be correctly represented by inversions. Also observations of the column-averaged dry air mole fraction of CO<sub>2</sub> from the Total Carbon Column Network (TCCON) may be available at Manaus, with other potential sites (e.g. Paramaribo, Suriname; [Belikov et al., 2017](#)) that could be installed in years to come, could also be exploited. These different types of data should not be analyzed as independent data streams but rather combined in an optimal way that accounts for their respective advantages. It would also be interesting, on a larger perspective, to extend my regional system to assimilate satellite data. Current satellite observations of the total atmospheric column of CO<sub>2</sub> (XCO<sub>2</sub>) from the GOSAT and OCO-2 satellites are limited by the dense cloud cover in Amazonia, which reduces the amount of available cloud-free soundings. But improvements to planned observing space-

borne instruments like GOSAT-2<sup>1</sup>—with updated optics and detectors—and OCO-3 should increase density of the measurements. The National Aeronautics and Space Administration (NASA) works on the concept of measuring CO<sub>2</sub> using laser technology in their Active Sensing of CO<sub>2</sub> Emissions over Nights, Days, and Seasons (ASCENDS<sup>2</sup>) instrument. The instrument should provide weekly, day- and nighttime measurements of XCO<sub>2</sub>, improving the picture about CO<sub>2</sub> fluxes. In December 2016 NASA announced an innovative space mission, the Geostationary Carbon Cycle Observatory (GeoCARB<sup>3</sup>), which should provide daily measurements of XCO<sub>2</sub> at ground resolutions of 5 – 10 km.

---

<sup>1</sup><http://www.gosat-2.nies.go.jp/about/mission>

<sup>2</sup><https://decadal.gsfc.nasa.gov/ascends.html>

<sup>3</sup><https://www.nasa.gov/press-release/nasa-announces-first-geostationary-vegetation-atmospheric-carbon-mission>

**A**

---

**Publication: On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia**



# On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over Amazonia

L. Molina<sup>1</sup>, G. Broquet<sup>1</sup>, P. Imbach<sup>2</sup>, F. Chevallier<sup>1</sup>, B. Poulter<sup>3</sup>, D. Bonal<sup>4</sup>, B. Burban<sup>7</sup>, M. Ramonet<sup>1</sup>, L. V. Gatti<sup>5</sup>, S. C. Wofsy<sup>6</sup>, J. W. Munger<sup>6</sup>, E. Dlugokencky<sup>8</sup>, and P. Ciais<sup>1</sup>

<sup>1</sup>Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ, IPSL, Gif-sur-Yvette, France

<sup>2</sup>Climate Change Program, Tropical Agricultural Research and Higher Education Center, Turrialba, Cartago 30501, Costa Rica

<sup>3</sup>Department of Ecology, Montana State University, Bozeman, MT 59717, USA

<sup>4</sup>INRA, UMR EEF, 54280 Champenoux, France

<sup>5</sup>CNEN – IPEN – Lab. Quimica Atmosferica, Av. Prof. Lineu Prestes, 2242, Cidade Universitaria, São Paulo, SP, Brazil

<sup>6</sup>Harvard University, School of Engineering and Applied Sciences, Department of Earth and Planetary Sciences, 20 Oxford Street, Cambridge, MA 02138, USA

<sup>7</sup>INRA, UMR Ecofog, Avenue de France, 97387 Kourou CEDEX, Guiana

<sup>8</sup>NOAA Earth System Research Laboratory, Global Monitoring Division, Boulder, CO 80305-3337, USA

Correspondence to: L. Molina (luis.molina@lscce.ipsl.fr)

Received: 1 September 2014 – Published in Atmos. Chem. Phys. Discuss.: 21 January 2015

Revised: 15 June 2015 – Accepted: 30 June 2015 – Published: 28 July 2015

**Abstract.** The exchanges of carbon, water and energy between the atmosphere and the Amazon basin have global implications for the current and future climate. Here, the global atmospheric inversion system of the Monitoring of Atmospheric Composition and Climate (MACC) service is used to study the seasonal and interannual variations of biogenic CO<sub>2</sub> fluxes in Amazonia during the period 2002–2010. The system assimilated surface measurements of atmospheric CO<sub>2</sub> mole fractions made at more than 100 sites over the globe into an atmospheric transport model. The present study adds measurements from four surface stations located in tropical South America, a region poorly covered by CO<sub>2</sub> observations. The estimates of net ecosystem exchange (NEE) optimized by the inversion are compared to an independent estimate of NEE upscaled from eddy-covariance flux measurements in Amazonia. They are also qualitatively evaluated against reports on the seasonal and interannual variations of the land sink in South America from the scientific literature. We attempt at assessing the impact on NEE of the strong droughts in 2005 and 2010 (due to severe and longer-than-usual dry seasons) and the extreme rainfall conditions registered in 2009. The spatial variations of the seasonal and interannual variability of optimized NEE are also investigated.

While the inversion supports the assumption of strong spatial heterogeneity of these variations, the results reveal critical limitations of the coarse-resolution transport model, the surface observation network in South America during the recent years and the present knowledge of modelling uncertainties in South America that prevent our inversion from capturing the seasonal patterns of fluxes across Amazonia. However, some patterns from the inversion seem consistent with the anomaly of moisture conditions in 2009.

## 1 Introduction

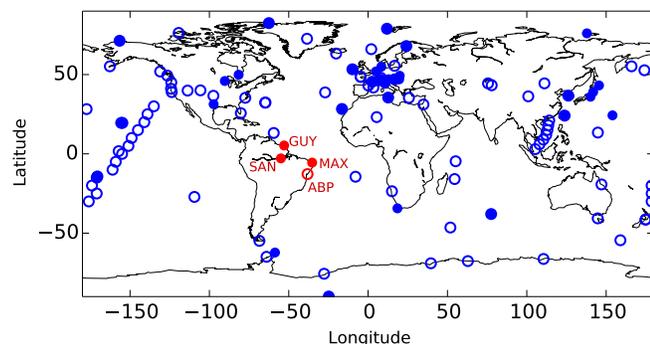
The forests of Amazonia cover 6.77 million km<sup>2</sup> (INPE, 2011). It is the world's largest continuous area of tropical forest and reservoir of aboveground organic carbon (Malhi et al., 2008). Changes in the carbon dynamics of this ecosystem thus have global significance (Wang et al., 2013). However, the natural variability of CO<sub>2</sub> exchange in Amazonia, as well as its short- and long-term response to natural and anthropogenic disturbance across scales, is still poorly understood and a topic of active research.

There is intense debate about the timing and magnitude of the seasonal cycle of CO<sub>2</sub> fluxes across Amazonia. Studies employing remote sensing data as a proxy for canopy photosynthetic activity have suggested a widespread enhancement of gross primary productivity of the Amazonian rainforest during the dry season (Huete et al., 2006). Yet direct and continuous measurements of net ecosystem exchange (NEE) between the atmosphere and forest canopy at a local scale (from 1 ha to 1 km<sup>2</sup> scale) based on eddy-covariance (EC) systems do not support such large-scale behaviour. Several EC observations in central eastern Amazonia (Saleska et al., 2003) and north-eastern Amazonia (Bonal et al., 2008) also indicate that tropical forest areas take up CO<sub>2</sub> during the dry season, but similar EC studies in central Amazonia have suggested an opposite seasonality (Grace et al., 1996; Araújo et al., 2002). Finally, remote sensing measurements of the vertically integrated columns of CO<sub>2</sub> (XCO<sub>2</sub>) retrieved from the GOSAT satellite suggest stronger CO<sub>2</sub> uptake during the wet season in southern Amazonian forest than during the dry season (Parazoo et al., 2013). These measurements thus reveal a large heterogeneity in space of the phase of the seasonal cycle of NEE within Amazonia. However, most dynamic global vegetation model (DGVM) simulations predict stronger uptake during the wet season throughout Amazonia (Verbeeck et al., 2011; Saleska et al., 2003; Baker et al., 2008; Poulter et al., 2009), although limitations related to mortality or land use restrict the ability of these generic global models to simulate CO<sub>2</sub> fluxes and carbon stocks of Amazonian forest (Gloor et al., 2012).

Uncertainty associated with potential spatial heterogeneity is also apparent in the estimates of the interannual variability (IAV) of CO<sub>2</sub> fluxes in Amazonia in particular during years with extreme climatic conditions. Remote sensing observations during the severe Amazonian drought of 2005 suggested a widespread enhancement of photosynthetic activity, or greening, across Amazonia (Saleska et al., 2007). The resilience of forests to water stress suggested by the “drier-yet-greener” papers was originally attributed to a combination of deep rooting, hydraulic redistribution and more available solar radiation (Saleska et al., 2007). However, the validity of enhanced vegetation index satellite data has been recently challenged by Morton et al. (2014) and by losses in canopy functioning detected in radar-based measurements (Saatchi et al., 2012). The observations from optical satellite sensors remain controversial because other studies did not find such an impact of droughts on Amazonian forest (Xu et al., 2011; Samanta et al., 2010, 2012). Moreover, observations of microwave backscatter from QuickSCAT have suggested large-scale persistent negative effects of the drought of 2005 on forest canopy structure (Saatchi et al., 2012). Biometry measurements, consisting of periodic measurements of the allocation of photosynthetic products to wood growth, provide another perspective on the effects of drought on Amazonian forest trees. In a large-scale, long-term biometric study, Phillips et al. (2009) found a rever-

sal of the carbon sink due to the effect of the drought of 2005 on tree mortality. This is consistent with a synthesis of yearly estimates of natural fluxes (NEE plus biomass-burning emissions) from an ensemble of DGVMs compiled at <http://www.globalcarbonatlas.org>.

The scientific community has used atmospheric inversions for more than 2 decades in an effort to improve the knowledge of CO<sub>2</sub> fluxes at a large scale. Whereas EC or biometric studies give flux estimates that are valid at the local scale (Ometto et al., 2005), atmospheric inversion offers the possibility to derive measurement-based estimates for the whole of Amazonia, with spatial resolutions larger than 500 km, provided that atmospheric observations can adequately sample the Amazonian flux signal. Inversions use available measurements of atmospheric CO<sub>2</sub> to provide corrections to prior surface flux estimates using an atmospheric transport model and statistical inversion methods. The method estimates statistically optimal fluxes within the boundaries of uncertainties in the measurements, the transport model and prior flux estimates (Enting et al., 1995; Ciais et al., 2010). The flux corrections spread beyond the vicinity of the measurement footprint, as defined by the transport model, through hypotheses on the spatial and temporal correlation of the uncertainties in the prior fluxes. We define, hereafter, the tropical South America (TSA) region as the continental land encompassed between 16.25° N–31.25° S and 84.38–28.18° W, which covers the whole Amazonian forest. Peylin et al. (2013) show that the different inverted seasonal cycles and IAVs of natural CO<sub>2</sub> fluxes from several state-of-the-art global atmospheric inversions are characterized by a large scatter over a very similar tropical area of South America. This is explained by the variety of prior estimates used by the different global inversion systems and by the large-scale corrections that are applied in regions poorly covered by observation networks, such as TSA, in order to balance the global CO<sub>2</sub> budget rather than to match local measurements. For these reasons, atmospheric inversions have not been included in the review of the carbon cycle in South America made by Gloor et al. (2012). Lloyd et al. (2007) and Gatti et al. (2010) applied the principle of atmospheric inversion to exploit vertical CO<sub>2</sub> profile data from airborne measurements in Amazonia. Their studies, based on measurements near Manaus in central Amazonia (Lloyd et al., 2007) and Santarém in eastern Amazonia (Gatti et al., 2010), constitute important efforts to constrain surface CO<sub>2</sub> fluxes at regional scale, measuring and exploiting some of the few atmospheric data sets available for South America. Their results suggested CO<sub>2</sub> efflux from the ecosystem during the wet season in eastern Amazonia. By analysing vertical CO<sub>2</sub> profiles collected approximately every 2 weeks over the period 2010–2011, the recent study of Gatti et al. (2014) provided a basin-scale picture that not only confirms this regional signal but also suggests an opposite pattern in southern and western Amazonia. Their study reported on the first data-driven estimate of CO<sub>2</sub> fluxes for the whole Amazon



**Figure 1.** Location of the surface stations used in this study. Blue indicates surface stations used in MACCv10.1; red shows the surface stations in South America added to the previous setup of MACCv10.1. Filled circles are stations with continuous measurements; open circles are sites with discrete air sampling.

basin and it provides insight into the sensitivity of this important ecosystem to moisture stress. It suggests the importance of conducting such estimates over longer time periods.

Our goal here is to study the seasonal cycle and IAV of NEE over Amazonia during 2002–2010. This period offers the opportunity to investigate significant anomalies in the interannual variability of carbon fluxes, particularly those associated with the severe droughts of 2005 and 2010 and the extreme rainfall registered across the Amazon basin in 2009 (Marengo et al., 2010). The study is based on the global Monitoring of Atmospheric Composition and Climate (MACC) inversion system initially described by Chevallier et al. (2010) (hereafter CH2010). We used version 10.1 of the MACC CO<sub>2</sub> inversion product released in August 2011. We also use a similar inversion in which we add four ground-based atmospheric measurement sites surrounding the northeast of Amazonia to the assimilated data (Fig. 1). Despite the limitations of the state-of-the-art global inversion approach in South America, highlighted above and by Gloor et al. (2012), our analysis of these MACC inversions can help characterise the temporal variations in the NEE over Amazonia for several reasons. First, it relies on a detailed evaluation of the inversion results over and within this region, hoping that some reliable inversion patterns can be isolated. Such a detailed evaluation has not been conducted in the above-mentioned intercomparisons of the global atmospheric inversions in TSA. It makes sense to conduct it here on the MACC inversions since the MACC system uses a variational inversion which solves for the fluxes at  $\sim 3^\circ$  and 8-day spatial and temporal resolution. Second, the use of the stations located in the region can strengthen the robustness of the inversion results through a significantly increased sampling of the atmospheric signature of the fluxes in Amazonia. In particular, we are the first to use continuous measurements from French Guiana. The assessment of the impact of these stations on the inverted NEE (based on the comparison between our dif-

ferent MACC inversions with and without these stations) can help identify the reliable patterns of the inversion.

The rest of this paper is structured as follows. We present each component of the standard MACCv10.1 inversion setup and the use of the additional sites around Amazonia in Sect. 2. The results of the inversions, with a focus on the impact of these additional sites, and their comparison to an independent flux estimate are presented in Sect. 3. In Sect. 4, we discuss the results and conclude the study.

## 2 The inversion method

This study builds on MACC, the global atmospheric inversion framework (whose first version is described in detail in CH2010), to correct a prior estimate of NEE from the model ORCHIDEE (Organizing Carbon and Hydrology in Dynamic Ecosystems, Krinner et al., 2005) and of ocean fluxes, based on the assimilation of in situ measurements of atmospheric CO<sub>2</sub> mole fractions into a global atmospheric transport model. The approach relies on a Bayesian framework to estimate the conditional probability of the “true” NEE and ocean fluxes given the statistical information from the prior fluxes and the set of in situ measurements of atmospheric CO<sub>2</sub> (hereafter *observations*). Assumption of unbiased Gaussian distribution of the uncertainties in the prior fluxes and of those underlying the simulation of the observations using the transport model allows us to derive an updated estimate of NEE and ocean fluxes (hereafter the posterior fluxes) that also has an unbiased Gaussian distribution. The statistically optimal fluxes (i.e. the mean of the posterior distribution of the fluxes) are found by calculating the minimum of the cost function (Tarantola, 2005):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y}^o - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y}^o - H(\mathbf{x})), \quad (1)$$

where  $\mathbf{x}$  is the control vector and mainly denotes the NEE (defined as the difference between the gross CO<sub>2</sub> uptake through photosynthesis and output through total ecosystem respiration) and air–ocean exchanges that are optimized at a chosen spatial and temporal resolution.  $\mathbf{x}^b$  represents the prior NEE and ocean fluxes, and  $\mathbf{y}^o$  is the vector of observations.  $H$  is the operator projecting  $\mathbf{x}$  into the observation space and is based on an atmospheric transport model and fossil fuel and biomass-burning CO<sub>2</sub> emission estimates.

$\mathbf{B}$  and  $\mathbf{R}$  are the covariance matrices of the normal distribution of the uncertainty in  $\mathbf{x}^b$  (the “prior uncertainty”) and of the sum in the observation space of the other uncertainties when comparing  $H(\mathbf{x}^b)$  to  $\mathbf{y}^o$  respectively (the “observation errors”). The latter includes the measurement, model transport and model representation errors. A complete solution to the inversion problem requires the estimation of the uncertainty in the optimized fluxes (the “posterior uncertainty”), which is a function of the prior and of the observation errors.

As explained below in Sect. 2.1, this estimation was not performed in this study. The following sections present a brief description of each component of the inversion configuration used in this study with a focus on parameters that are specific to this study, while CH2010 provides more details on the parameters which apply to all the MACC inversion configurations.

## 2.1 Inversion modelling setup

The link between CO<sub>2</sub> fluxes and observations in the MACC inversion is simulated by the global circulation model of the Laboratoire de Météorologie Dynamique (LMDZ) (version 4, Hourdin et al., 2006), which is the atmospheric component of the coupled climate model of the Institut Pierre Simon Laplace (IPSL-CM4). Tracer transport is simulated by LMDZ at a horizontal resolution of  $3.75^\circ \times 2.75^\circ$  (longitude  $\times$  latitude) and with a vertical resolution of 19 levels between the surface and the top of the atmosphere. LMDZ is nudged to winds modelled by the European Centre for Medium-Range Weather Forecasts (ECMWF). Prior NEE in MACCv10.1 was estimated at  $3.75^\circ \times 2.75^\circ$  and 3 h resolution from a global simulation of the ORCHIDEE model at  $0.7^\circ$  resolution by Maignan et al. (2011). ORCHIDEE was forced with the atmospheric conditions of ECMWF reanalysis ERA-Interim (Berrisford et al., 2009). The ORCHIDEE NEE did not take into account disturbance from land use or wildfires. Prior ocean–atmosphere CO<sub>2</sub> exchanges were obtained from the climatology of air–ocean CO<sub>2</sub> partial pressure difference by Takahashi et al. (2009).

To complement these fluxes that were controlled by the inversion, the  $H$  operator also included fixed estimates of the fossil fuel and biomass-burning CO<sub>2</sub> emissions. Fossil fuel emissions were obtained from the EDGAR-3.2 Fast Track 2000 database (Olivier and Berdowski, 2001), scaled annually with the global totals of the Carbon Dioxide Information Analysis Center. CO<sub>2</sub> emissions from biomass burning were taken from the Global Fires Emission Database version 2 (GFEDv2, Randerson et al., 2007). Assuming that the vegetation recovers rapidly from fire events, the CO<sub>2</sub> emissions from fires that affected the vegetation in a given year were offset by an equivalent compensatory regrowth CO<sub>2</sub> uptake evenly distributed throughout the year.

The inversion controlled 8-day mean daytime and nighttime NEE and 8-day mean ocean fluxes at the spatial resolution of the transport model. The analysis in this study focuses on NEE and thus the impact of the inversion on ocean fluxes is not detailed here, but Sect. 3.2 still uses an illustration of this impact to raise insights into the corrections from the inversion over land. At the grid scale, uncertainties in the prior NEE are estimated to be proportional to the heterotrophic respiration fluxes from ORCHIDEE. Spatial correlations of the uncertainties in  $\mathbf{B}$  decay exponentially as a function of the distance between corresponding pixel-based estimates of the fluxes with a length scale of 500 km for NEE (1000 km

for ocean fluxes). Temporal correlations of the uncertainties decay exponentially as a function of the lag time between the corresponding 8-day mean daytime or nighttime estimate of the fluxes with a timescale of 1 month but without correlation between daytime and nighttime uncertainties. The resulting correlations in  $\mathbf{B}$  are estimated as the product between the temporal and the spatial correlations. This setup of the correlations for  $\mathbf{B}$  is based on the estimates by Chevallier et al. (2006) and Chevallier et al. (2012) of differences between the NEE simulated by ORCHIDEE and EC flux measurements (mostly located in the Northern Hemisphere).

In the inversion framework, the misfits between simulated CO<sub>2</sub> mole fractions and the measurements that are not due to uncertainty in the prior NEE or ocean fluxes must be accounted for in the covariance matrix  $\mathbf{R}$ . Uncertainties in fire and anthropogenic CO<sub>2</sub> emissions are assumed to have negligible impact at the measurement locations used here. Therefore, they are ignored in the setup of  $\mathbf{R}$ . Following CH2010, the measurement errors are assumed to be negligible in comparison to the uncertainties in the transport model. Model transport and representation errors are modelled as half the variance of the high-frequency variability of the deseasonalised and detrended CO<sub>2</sub> time series of the measurements that are assimilated at a given station. The resulting values of these model errors for the stations in South America will be discussed in Sect. 3.1.

There is a moderate confidence in the adequacy of these error statistics assigned in the global inversion system for the specific TSA area studied here, both because  $\mathbf{B}$  was designed mostly with statistics gathered in the Northern Hemisphere and because  $\mathbf{R}$  may not well account for the uncertainty in the atmospheric convection model, while this could be high in Amazonia (Parazoo et al., 2008). We also investigate here variations of the fluxes within TSA at spatial scales that are not much larger than the e-folding correlation length in  $\mathbf{B}$ , and these variations in the inversion results may be affected by our simple hypothesis of isotropic correlations in the prior uncertainty. This lack of confidence in the input error statistics weakens our confidence in the posterior error statistics that can be derived based on the inversion system, even though they may be realistic at zonal scale for the tropics (Chevallier and O'Dell, 2013). In this context, and given the relatively high computational burden of the posterior uncertainty computations for grid-point inversion systems (using Monte Carlo approaches with ensembles of inversions, Chevallier et al., 2007), we do not derive these posterior uncertainties for our domain and its sub-domains.

However, we will see at the beginning of Sect. 3 that the inverted fluxes are more consistent with the CO<sub>2</sub> atmospheric observations in TSA than the prior fluxes and that their difference to the prior fluxes over TSA (i.e. the flux increments generated by the inversion in order to better fit with the observations) are significant. This indicates that the inverted fluxes are strongly driven by the atmospheric data and as such are

Station	Period of data availability for this study										Principal investigator
	2002	2003	2004	2005	2006	2007	2008	2009	2010		
Arembepe (ABP)											E. Dlugokencky ESRL, NOAA, Boulder, Colorado, USA
Guyaflux (GUY)											D. Bonal, INRA, Nancy, France and B. Burban, INRA, Kourou, French Guiana
Maxaranguape (MAX)											B. Munger Harvard University, Cambridge, Massachusetts, USA
Santarém (SAN)											S. Wofsy Harvard University, Cambridge, Massachusetts, USA

**Figure 2.** List of surface stations over South America added to the previous setup in MACCv10.1.

worth analysing. This also suggests that the inversions yield a large uncertainty reduction for TSA.

## 2.2 Assimilated data

MACCv10.1 assimilated measurements of atmospheric CO<sub>2</sub>, expressed as dry air mole fractions in  $\mu\text{mol mol}^{-1}$  (abbreviated ppm) from 128 surface sites: 35 continuous measurement stations and 93 sites with measurements of CO<sub>2</sub> from discrete air samples collected approximately weekly. Twenty-nine sites are located in the tropics, but only two had continuous measurements over the analysis period and none of them were in TSA. In a similar inversion conducted specifically for this study, called INVSAM hereafter, we added new data from four surface sites located in the TSA region. Figure 1 shows the measurement sites used by MACCv10.1 and the four stations added in INVSAM. In the following of this section, we focus on the description of these four stations and on the selection and representation of their data. Details on the data selection and representation at the sites used by MACCv10.1 are provided in CH2010.

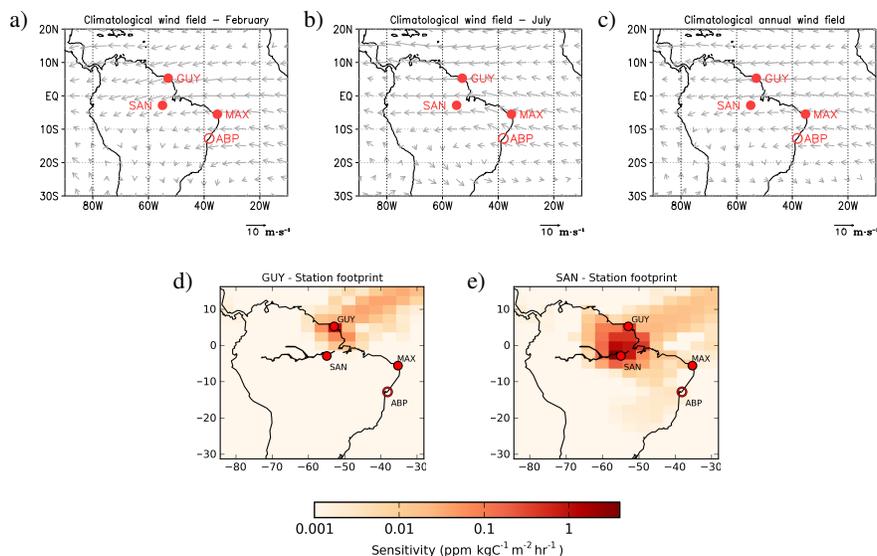
Arembepe (ABP) (12.77° S, 38.17° W; 1 m a.s.l.) and Maxaranguape (MAX) (5.51° S, 35.26° W; 15 m a.s.l.) are coastal stations. The ABP site is located at the edge of the beach, where vegetation consists mostly of grass and beach plants. Data were collected at approximately 3 m above the ground and consisted of weekly measurements of atmospheric CO<sub>2</sub> with discrete air samples, specifically under on-shore wind conditions when wind speed  $> 2 \text{ m s}^{-1}$ . Air samples were collected preferentially during the afternoon to avoid the influence of recycled air transported from land to the ocean by land breeze during the night and early morning and transported back to land by sea breeze during the morning. The MAX site is located on a cliff right next to the coast and is surrounded by grass and beach plants. At MAX, CO<sub>2</sub> was measured with a continuous analyzer at approximately 3 m above the ground, and data were reported as 30 min averages. This site is strongly under marine influence: winds are in general  $> 10 \text{ m s}^{-1}$ , and wind direction varies preferentially between 100° and 140° (Kirchhoff et al., 2003) at its location, so that the measurements were taken mostly under on-shore wind conditions. Wind and CO<sub>2</sub> measurements at MAX indicate high CO<sub>2</sub> variations when the wind comes

from land. These variations may be strongly influenced by the emissions from the nearby city of Maxaranguape (Kirchhoff et al., 2003). However, as in ABP, this does not occur during the afternoon, when the wind conditions are dominated by sea breeze (Law et al., 2010).

The Guyaflux site (GUY) (5.28° N, 52.91° W; 40 m a.s.l.) is located at approximately 11 km from the coast and is surrounded by undisturbed tropical forest. At GUY, measurements were taken at approximately 55 m above the ground (Bonal et al., 2008). They were made with a continuous analyser, and data were reported as hourly averages. The Santarém site (SAN) (2.85° S, 54.95° W; 78 m a.s.l.) is located in the tropical Tapajós National Forest, near km 67 of the Santarém–Cuiabá highway, at approximately 750 km from the coast. Measurements were made at eight vertical levels ranging from  $\sim 1$  to  $\sim 62$  m above the ground with continuous analyzers, but only data from the highest level were used in INVSAM. Data were reported as hourly averages.

Figure 2 illustrates the temporal coverage of the observations available in TSA during the simulated period (2002–2010). There is little overlap among the site records due to calibration problems, interruption of the measurements (e.g. at MAX) and the fact that some stations have been installed only recently (e.g. at GUY). The longest records were from ABP (3 years: 2007–2009) and SAN (4 years: 2002–2005). Data from the four new sites in TSA have been calibrated on the WMO-X2007 CO<sub>2</sub> scale managed by the ESRL/NOAA.

Prevailing winds in the lower troposphere across TSA convey air masses entering from the Atlantic Ocean near the Equator, across the continent and back into the southern Atlantic Ocean generally south of 20° S. There are no critical seasonal variations of the mean winds in the area so that this typical behaviour applies throughout the year. The climatology of wind fields from the NCEP/NCAR reanalysis (over the period 1981–2010) for February, July and annual mean, shown in Fig. 3, illustrates this typical circulation pattern. This confirms that the variations of CO<sub>2</sub> at coastal stations (ABP, MAX) are mainly influenced by air–ocean exchanges and fluxes in distant lands. These stations should thus provide more information on the atmospheric CO<sub>2</sub> content upwind of TSA than on the fluxes within Amazonia. Figure 3 also shows that GUY and SAN receive a signal from the ecosys-



**Figure 3.** Top: location of assimilated surface stations in South America and climatological wind speed/direction for February (a), July (b) and annual mean (c), averaged over 1981–2010 between the surface and a level of 600 hPa (source: NCEP/NCAR reanalysis). Sensitivity of surface atmospheric CO<sub>2</sub> mole fractions measured on 20 February 2009 at 10:00 UTC, at Guyaflux (07:00 LT) (d) and Santarém (06:00 LT) (e), to a constant increment of surface fluxes during the 2 days prior to the measurement. Sensitivity values are expressed in log scale. Open circles: sites with discrete air samplings. Filled circles are measurements taken with continuous analysers.

tems of the north-eastern Amazon basin. Despite GUY being not far from the coast considering the Amazon-wide scale, this site is still located inland, in an area covered by undisturbed tropical wet forest. SAN is located considerably further inland than GUY. Typical influence functions of fluxes for observations at GUY and SAN (the observation “footprints” in Fig. 3b and c respectively) illustrate that the sensitivity of instantaneous mole fractions to the fluxes rapidly decreases with the distance mainly due to the typically moderate horizontal wind speeds, so that they should bear a strong signature of local fluxes, i.e. of the NEE in north-eastern Amazonia. This and the fact that the geographical distance between the sites in the TSA region ranges from 1000 to 2600 km, i.e. up to 5 times the correlation length scale in the matrix **B**, could suggest that the area well constrained by the sites in the TSA region through inversion is limited. However, as illustrated in Fig. 3, the station footprints also have modest values over very extensive areas, which may also result in significant large-scale constraint from the inversion on the land flux estimates. This will be analysed below in Sect. 3.2.

We assimilated observations from the South American sites between 12:00 and 15:00 local time, when the boundary layer is well developed and likely to be well represented by the transport model (Butler et al., 2010; Gatti et al., 2010). Such a selection of the afternoon data results in ignoring the measurements under off-shore flow at MAX and thus the potential for capturing a clear signature of the regional NEE at this site such as at ABP. However, this potential is rather low since under off-shore flow conditions the signal at MAX

is also connected to the local anthropogenic emissions, and the inversion cannot reliably exploit such a signature of the regional NEE when the dynamics of the planetary boundary layer are poorly represented by the atmospheric transport model. Observations were also screened for low wind speed ( $> 2 \text{ m s}^{-1}$ ), thus removing the effect of local emissions (and sinks) that may not be well captured by the transport model at resolution  $3.75^\circ \times 2.5^\circ$ . Under such on-shore flow conditions, the model correctly simulates CO<sub>2</sub> in the grid cells corresponding to the horizontal location of the coastal sites, even though these grid cells bear a significant NEE due to the overlapping of both land and ocean. This reduces the need for ad hoc changes of the model grid cells to better represent CO<sub>2</sub> at the coastal sites (e.g. Law et al., 2010). In a general way, we choose to represent the four measurements sites using the model horizontal grid cell in which they are located since, for each site, it yields better statistical fit between the prior simulations and the selected measurements than when using neighbour grid cells.

### 2.3 Analysis of an alternative estimate of the NEE for the evaluation of the inversions

Our analysis of the inversion results is compared to the independently derived NEE estimated by Jung et al. (2011) (hereafter J2011). J2011 used model tree ensembles (MTE), a machine-learning technique, to upscale FLUXNET eddy-covariance observations, based on remote sensing, climate and land-use data as drivers, thereby producing gridded estimates of NEE and other surface fluxes at the global scale

at 0.5° resolution. As discussed in J2011, large uncertainties affect their annual mean NEE estimates and associated seasonal and interannual variations. This is likely particularly true in TSA region, where few FLUXNET measurements are available. Yet its comparison to the NEE from the inversion could give useful insights for the analysis of the latter.

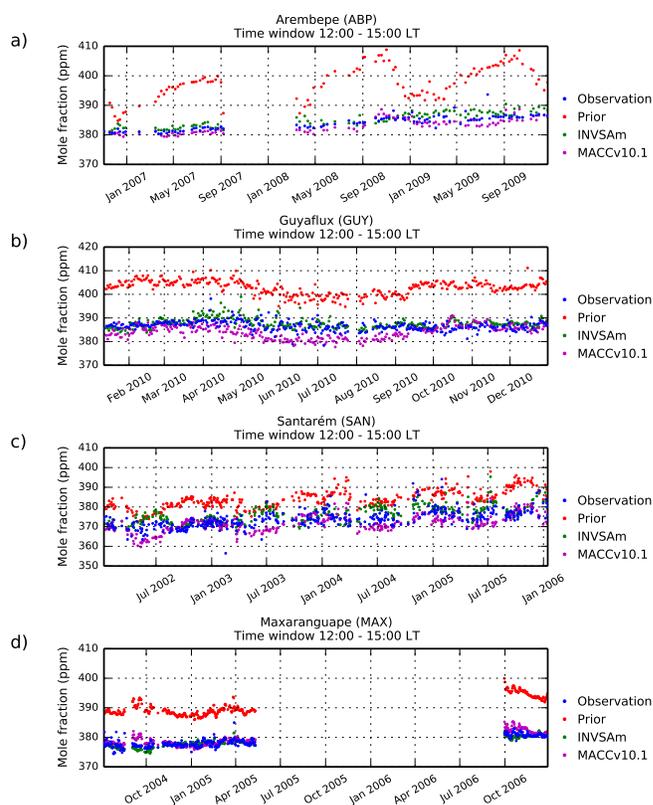
### 3 Results

In this section we first analyse the statistical misfits between observations and simulated mole fractions from prior and posterior fluxes at the sites in the TSA area, as a measure of the efficiency of the inversion in reducing the misfits to the measurements. This is a first indicator of the significance of the corrections applied to the fluxes. We then examine the amplitude and spatial distribution of the increments from both inversions to give a further indicator of this significance and to characterise the impact of assimilating the measurements from the sites in South America. Finally we focus on the impact of the inversions on the seasonal patterns and IAV of NEE which are the aim of this study. This analysis is supported by the comparison to the product of J2011.

#### 3.1 Comparison to observed CO<sub>2</sub> mole fractions

The time series of assimilated observations and the corresponding simulated CO<sub>2</sub> mole fractions using the prior fluxes, the inverted fluxes from MACCv10.1 and that from INVSAM at the four sites in the TSA region are plotted in Fig. 4. The statistics of the misfits between these measured and simulated CO<sub>2</sub> mole fractions are summarised in Fig. 5. At each site in the TSA region, the smallest quadratic mean and standard deviation of the misfits between the simulations and the observations were obtained with INVSAM, which is a logical consequence of the assimilation of these observations. However, the misfits are also strongly decreased at all sites when comparing MACCv10.1 to the prior simulation. While, compared to the prior simulation, MACCv10.1 strongly decreases the standard deviation of the misfits at MAX and ABP, it does not significantly reduce it at GUY and SAN. The decrease of the misfits at all sites in MACCv10.1 is thus explained by the strong decrease of the bias in these misfits. Indeed, both inversions critically reduce a large-scale bias over TSA, since the presence of a few marine stations on the globe is enough to introduce this effect by correcting the global growth rate of CO<sub>2</sub> (CH2010). However, the information from the local network significantly impacted the seasonality of the simulated CO<sub>2</sub> in the TSA region.

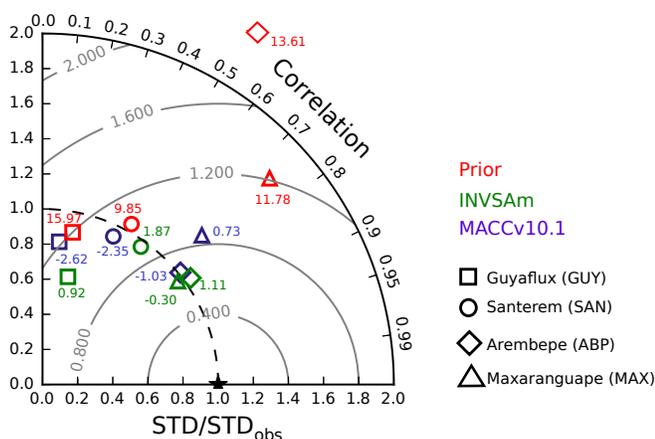
The resulting optimized mole fractions from INVSAM generally shifted from a minimum to a maximum around June every year at SAN or from a maximum to a minimum around October (both in 2004 and 2006) at MAX with respect to the prior simulation and MACCv10.1 (Fig. 4c) and in agreement with the observations. While yielding a phase



**Figure 4.** Comparison of assimilated CO<sub>2</sub> observations (blue) and corresponding simulated mole fractions using prior fluxes (red), INVSAM (green) and MACCv10.1 (purple). Measurements were collected at Arembepe (a), Guyaflux (b), Santarém (c) and Maxaranguape (d). Data shown here correspond to daily average mole fractions between 12:00 and 15:00 LT, when wind speed > 2 m s<sup>-1</sup>. Note that the timescale differs between plots.

of seasonality at GUY comparable to that of the prior simulation and MACCv10.1 and comparable to that of the data, INVSAM exhibits a significant rescaling of the seasonal variations in the period from May to September at this site (Fig. 4b) compared to these two other simulations, in agreement with the observations. At SAN, during the austral fall–winter, while the misfits are negative with MACCv10 they become positive with INVSAM. The positive increments from the assimilation of data at SAN (no other data are assimilated in TSA in 2002 and 2003) are thus too high.

Subsequently, when compared to MACCv10.1, INVSAM improves the amplitude of the seasonal variations of the simulated mole fractions with respect to the prior simulation at GUY and MAX and does not impact it at SAN. At ABP, the seasonality is less visible in both the measurements and the inversion posterior simulations and it is difficult to assess whether INVSAM improves it compared to MACCv10.1, but both inversions dramatically decrease the large amplitude of the prior seasonal variations, consistent with the data. The best correlations with the observations are obtained with IN-

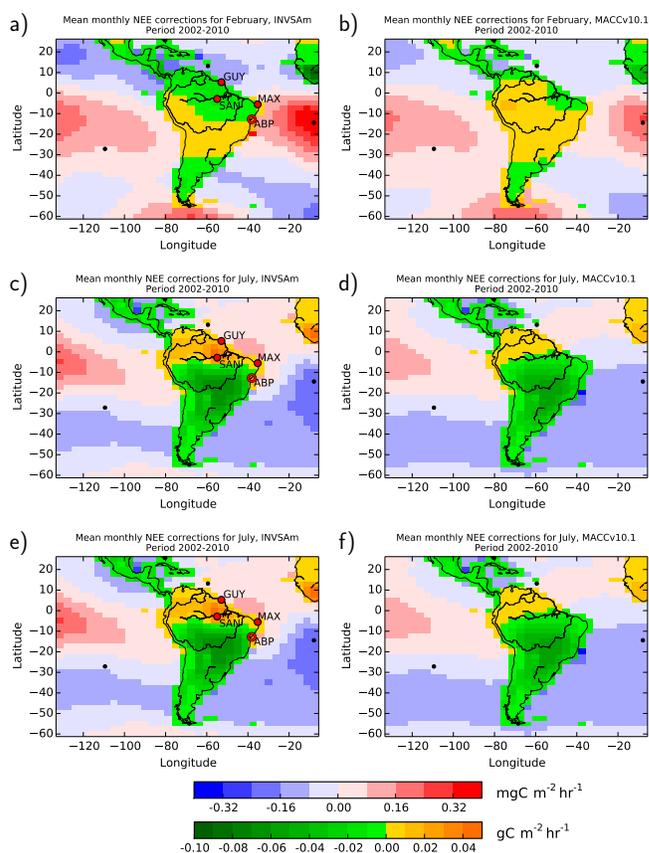


**Figure 5.** Taylor diagram of the statistics of misfits between observations and simulated CO<sub>2</sub> mole fractions between 12:00 and 15:00LT at Guyaflux (square), Santarém (circle), Arembepé (diamond) and Maxaranguape (triangle), when wind speed > 2 m s<sup>-1</sup>, using prior fluxes (red), INVSAm (green) and MACCv10.1 (purple). Radial distance from the origin: ratio of SD of simulated mole fractions and SD of the observations. Angle measured from the y axis: coefficient of correlation. Numbers next to the symbols: bias (in ppm). Grey circles: SD of the misfits (in ppm).

VSA<sub>m</sub> at all sites (Fig. 5). The values of these correlations remained generally low, ranging from 0.23 at GUY to 0.81 at ABP. These correlations are based on comparison of daily CO<sub>2</sub> mole fractions while the inversions control 8-day mean fluxes, which strongly limits the ability to impact the mole fractions at higher temporal resolution and can thus explain the low correlation values. Correlations between time series of observed and simulated monthly mean mole fractions are higher than those for daily values, ranging from 0.76 at GUY to 0.92 at ABP for INVSAm, with which, again, these correlations are the highest.

The significance of the reduction of the misfits between the mole fractions observed and simulated from the inversion is seen from the comparison between the standard deviations of these misfits and the estimate of the standard deviation of the observation errors (i.e. of the transport model errors) for hourly values in the configuration of the **R** matrix (Table A1 in the Supplement). According to this comparison, the prior misfits are much larger than the observation errors at ABP, MAX and GUY but are slightly smaller than these at SAN. Misfits between MACCv10.1 and the observations are similar to the prior misfits at SAN and GUY and are much smaller than the prior misfits (and smaller than the 95 % confidence interval of the observations) at the coastal ABP and MAX sites. Misfits are further decreased when assimilating the data from the South American sites: they are about the standard deviation of the observation errors at all sites but GUY (where they are twice as large).

These results suggest that the assimilation of data in the TSA region helped improve the phasing of the seasonal vari-



**Figure 6.** Spatial distribution of 2002–2010 mean flux corrections at the transport model resolution ( $3.75^\circ \times 2.50^\circ$ ) to ORCHIDEE from INVSAm (left) and MACCv10.1 (right) over an area larger than TSA region: mean for February (a, d), July (b, e) and the full period 2002–2010 (c, f). Flux increments over land and ocean are represented with two distinct colour scales and units: green–yellow for land, in  $\text{gC m}^{-2} \text{h}^{-1}$ ; blue–red for ocean, in  $\text{mgC m}^{-2} \text{h}^{-1}$ . Red symbols are surface stations in South America added to the previous setup of MACCv10.1, where filled circles indicate locations of sites with continuous measurements; open circles indicate locations of sites with discrete air sampling. Black symbols are the surface stations used in MACCv10.1.

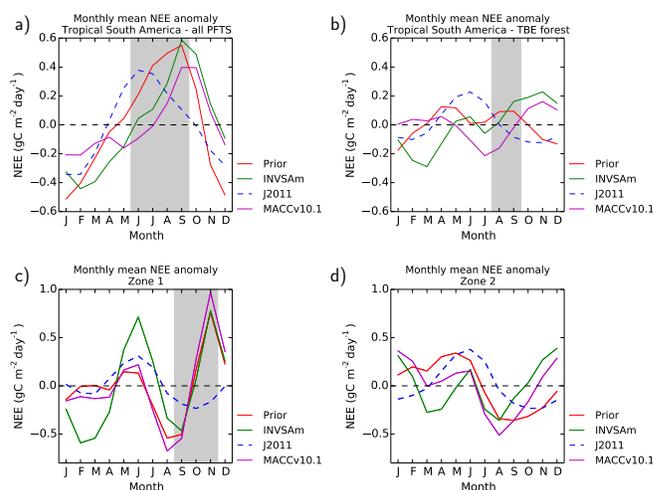
ations, whereas MACCv10.1 did not impact it. MACCv10.1 mainly improved the amplitude of the seasonal variations at the coastal sites and decreased the biases. INVSAm improved the amplitude of the seasonal variations at GUY. More generally, unlike MACCv10.1, INVSAm led to an improvement of the variability of the simulated CO<sub>2</sub> at the inland sites, which are more sensitive to the NEE in Amazonia.

### 3.2 Characterisation of the monthly to annual mean inversion increments to the prior fluxes

Figure 6 shows the spatial distribution of the mean corrections applied during the period 2002–2010 by INVSAm and MACCv10.1 over land and ocean, across an area that covers the TSA area and neighbour regions. Complementary to

this, Fig. S1 shows the spatial distribution of the corrections over land in the TSA region for the full 2002–2010 period and for the 2002–2005 and 2006–2010 sub-periods. Both give results for the full years and for the months of February and July. As such, these figures are indicative of the amplitude and spatial extent of the corrections from the inversions and of the impact of the assimilation of the measurements in South America. Figure S1 even dissociates the impact of assimilating data at SAN and MAX and that of assimilating data at MAX, ABP and GUY by splitting the results between the time periods when these two different sets of data are available. The analysis of the annual mean corrections and of mean corrections for February and July should also give first insights on the significance of the corrections applied to the seasonal cycle and IAV of the NEE in the TSA region.

Figure 6 depicts the increments from both inversions, showing large patterns which are nearly zonal (or along the prevailing winds) and overlap continuously over land and ocean. Since there is no correlation between the uncertainty in ocean and land fluxes in the **B** matrix, and given the typical length scale of the correlations in this matrix, this can be directly connected to the signature of atmospheric transport. The contiguous zonal patterns have alternate negative and positive flux increments. There is thus an opposition between corrections in the north and in the south of the TSA region. These corrections are rather negative in the north and positive in the south (positive in the north and negative in the south) during the austral summer (winter). As these corrections are stronger during the austral winter, it results in positive (negative) corrections in the north (south) at the annual scale. Such dipoles are a typical behaviour of inverse modelling systems in data-poor regions (Peylin et al., 2002). However, changes in the amplitude and latitudinal position of this zonal dipole appear to be the main impact from the assimilation of data in the TSA region. This dipole structure may thus yield sensible corrections to the NEE in the TSA area. The dipole has a high amplitude for MACCv10.1 and even higher for INVSAm. The increments from INVSAm to the annual fluxes often exceed 150 % of the prior estimate in terms of absolute values. The highest increments are obtained during austral winter and when the SAN data are available (during the period 2002–2005, see Fig. S1), which is in line with the fact that this site is located more inland than the others. Such high control of the data in the TSA region (even when checking the SAN and MAX or the MAX, ABP and GUY data sets only) over the zonal patterns of flux corrections also highlights the very large-extent impact of these data, and of the data in the Southern Hemisphere in general, despite the relatively small spatial correlation length scales in the **B** matrix and the limited area in which the station footprints are very high. The inversion also generates patterns of corrections of smaller spatial scale close to the measurement sites in the TSA region when these sites are used by the inversion. This raises hope that the NEE over the whole TSA region is strongly constrained by the observations but can also



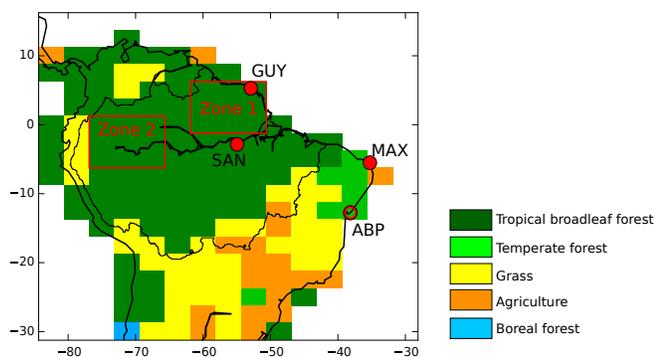
**Figure 7.** Monthly mean NEE anomaly integrated over (a) the TSA region and (b) over pixels dominated by TBE forests in ORCHIDEE for 2002–2010. The shaded areas denote dry seasons, defined as months with precipitation < 100 mm, based on monthly totals from TRMM data over 2002–2010. Estimates from prior fluxes (red), INVSAm (green), CH2010 (purple) and J2011 (dashed blue). (c–d) Monthly mean NEE integrated over the zones 1 (c) and 2 (d) that are defined in Fig. 8.

raise questions regarding the spatial variations of the corrections applied by the inversion to the NEE within the TSA region, at least when considering areas at more than 500 km from the measurement sites. However, various pieces of evidence (Figs. 5 and 6, the analysis of the decrease in misfits to the observations from the inversion in Sect. 3.1 and the previous analysis of the high increments to the monthly mean and annual mean NEE over the entire TSA region) indicate that the corrections from the inversion are significant.

### 3.3 Diagnostics of the biogenic CO<sub>2</sub> fluxes

#### 3.3.1 Seasonality

Figure 7a illustrates the mean seasonal cycle of NEE from the prior fluxes, J2011, MACCv10.1 and INVSAm over TSA. The mean for the full period 2002–2010 was removed because uncertainties in the long-term mean can be large for the inversions as well as for the J2011 product and because this long-term mean can differ significantly between the different estimates. Removing the mean allows us to focus on the seasonal variations. Hereafter, positive values of NEE indicate anomalous CO<sub>2</sub> release to the atmosphere; negative values indicate anomalous uptake by the ecosystems. The shaded area indicates the dry season, defined as months with precipitation < 100 mm according to data from the Tropical Rainfall Measuring Mission (TRMM 3B43 (v6) product), averaged over January 2002 to June 2010. The results of Fig. 7a are calculated considering all the plant functional types (PFTs) represented in ORCHIDEE over the TSA region. The vegeta-



**Figure 8.** Dominant PFTs for each transport model grid cell (i.e.  $3.75^\circ \times 2.50^\circ$ ) according to the ORCHIDEE vegetation map over the study region. Open circles show location of sites with discrete air sampling; filled circles show location of sites with continuous measurements. Zones 1 and 2 indicate areas for which the NEE is presented in Fig. 7c and d respectively.

tion map of ORCHIDEE, originally at a spatial resolution of  $0.72^\circ$ , was aggregated according to the transport model grid, and Fig. 8 illustrates the dominant PFTs in terms of area for each transport model grid cell.

Both the prior simulation and the inversions predict a maximum of NEE (i.e. likely a maximum of CO<sub>2</sub> release) in the dry season and a minimum of NEE (i.e. likely a maximum of CO<sub>2</sub> uptake) in the wet season (Fig. 7a). This behaviour is also seen in J2011. However, J2011 place the maximum of NEE during the transition between the wet and dry season while the prior simulation and the inversions place it at the end of the dry season. Even though the inversions seem to delay or lengthen this maximum, such a modification is not significant and their seasonal phasing is likely strongly constrained by the patterns of the prior fluxes. In particular, according to the comparison between INVSAM and MACCv10.1, the assimilation of data from the four stations in the TSA region does not seem to impact this phasing.

The inland data are prone to bear a stronger signature from fluxes in tropical broadleaf evergreen and raingreen (TBE) forests (Fig. 8), while the mean seasonal behaviour over the whole TSA region could be mainly related to other PFTs. Therefore, we isolate the results for the area of TBE forests, this area being defined by the selection the model grid cells dominated by this vegetation type. The configuration of the prior uncertainties in the inversion does not account for PFTs, so that the spread of the flux corrections in the inversions is not forced a priori to depend on vegetation type. We still expect that the variations in the measurements, when their footprint covers different distributions of PFTs, reflect differences in NEE of the PFTs. Consequently, the spatial patterns of the increments from the inversion may be consistent with the spatial patterns of NEE induced by the distribution of the different vegetation types. The mean seasonal cycle of NEE for the area of TBE forests within the TSA region is given in

Fig. 7b). The restriction of the analysis to the TBE forest does not show any clear correlation between NEE extremes and the phasing of wet and dry seasons neither when considering the NEE from the prior nor when considering the NEE from both inversion estimates. This is different from J2011, who indicate a maximum of the NEE a few months before the beginning of the dry season and a minimum of the NEE at the beginning of the wet season. The prior and the inversions indicate several local extremes of NEE throughout the year that may reflect the overlapping of significantly different seasonal cycles for different sub-regions within TBE forests.

The strong spatial heterogeneity of the time variations of the NEE in TBE forests has been discussed in the introduction. Figure S2 illustrates it this with results of local NEE mean seasonal cycle estimated from EC measurements across TSA. This figure also shows the mean seasonal cycle of the precipitation at these sites to illustrate the spatial heterogeneity of the drivers of NEE within TSA.

To examine whether the inversion captures this spatial variability of the fluxes, we analyse the seasonal variations of the NEE estimates for the two zones depicted in Fig. 8. Zone 1 was located in north-eastern Amazonia, close to the measurement stations SAN and GUY. Zone 2 was located in central eastern Amazonia. Both zones are mainly covered by TBE forests, according to the vegetation classification of ORCHIDEE. According to Malhi et al. (2009), eastern Amazonia is drier and shows a stronger seasonality than western Amazonia. However, we do not identify a clear pattern of NEE seasonal variations that could be driven by the rainfall seasonality in any of the two sub-regions, except for J2011 in Zone 1 (Fig. 7c), since the other estimates again exhibited maxima and minima of NEE during both dry and wet seasons. Actually, in Zone 2 (Fig. 7d) the dry season cannot be clearly identified. In this zone, the prior flux and the inversions indicated several maxima and minima of NEE, but J2011 exhibit, again, a clear seasonal cycle with a maximum in June and a minimum October as in Zone 1. While J2011 showed nearly the same amplitude and phasing of monthly mean NEE variations in both zones and over TBE forests (Fig. 7b), prior and inversions estimates of the seasonal variations differed both in phasing and amplitude among zones 1, 2 and the whole TBE forest area.

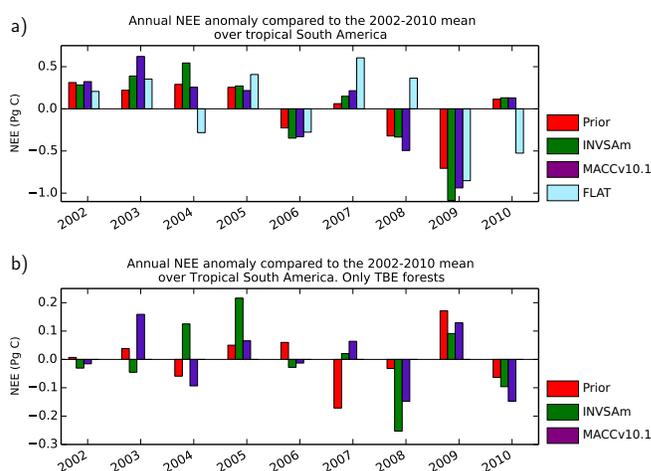
Divergent patterns are found in INVSAM with respect to MACCv10.1, which remains closer to the prior fluxes, even though the departure of MACCv10.1 from the prior NEE is significant in Zone 2 and for the whole TBE area (Fig. 7b and d). The comparison of these inversion results shows that significant flux corrections due to the assimilation of data in South America are applied in Zone 1 (Fig. 7c), i.e. in north-eastern Amazonia, where stations SAN and GUY are located. The influence of SAN over this zone is clearer when splitting the analysis period of the mean seasonal cycles between 2002–2005 and 2006–2010 (not shown). The differences between INVSAM and MACCv10.1 are more accentuated during the period 2002–2005 when SAN is active. However,

there are still significant changes between these two estimates during 2006–2010. The changes between MACCv10.1 and INVSAm in Zone 2 (Fig. 7d) are also significant, even though Zone 2 seems hardly observed by the TSA observation network. As analysed in Sect. 3.2, the control of the long-range dipole (of its amplitude and latitudinal position) by the measurements in region TSA explains such an impact of these measurements on the results in Zone 2, as well as that of measurements outside South America, which explains the departure of MACCv10.1 from the prior NEE in Zone 2. Zone 2 is actually located close to the frontier between the northern and southern patterns of the dipole in the TSA region. A latitudinal shift of the frontier through the assimilation of data in north-eastern Amazonia can thus easily imply that positive (negative) increments from the inversion are reverted into negative (positive) increments.

In an attempt at getting clearer seasonal patterns in some of the other sub-regions of Amazonia, two additional zones have been analysed, located in south-western and south-eastern Amazonia, where the dry season is potentially earlier and more extreme (Fig. S2a, d). Both sub-regions encompass areas where the impact of the droughts of 2005 and 2010 was the highest according to Lewis et al. (2011). The results, however, do not provide any further information than Fig. 7c and d and are not shown. J2011 still exhibit the same amplitude of the seasonal cycle and the same location of maximum and minimum NEE as in zones 1 and 2 despite the extent of the dry season. Prior fluxes and inversions still showed maxima and minima during the dry season in some cases, and the inversions introduce only slight modifications to the amplitude and phasing of the NEE relative to the prior simulation. This is an expected result due to insufficient data in the southern part of TSA to constrain fluxes in that region.

### 3.3.2 Interannual variability

Figure 9a depicts the annual NEE anomalies of the prior simulation, MACCv10.1, INVSAm and an additional inversion called FLAT, compared to their mean NEE over 2002–2010, aggregated over the whole TSA region (considering all PFTs). FLAT corresponds to a new inversion using, as a prior estimate, a “flat prior” whose annual anomalies are null over the TSA region. Using the standard prior NEE as a basis, the flat prior is built by offsetting the annual budgets of NEE over the TSA region so that they equal the mean annual NEE over TSA and over the 2002–2010 period from the standard prior NEE. The spatial variability and the temporal variability at scales smaller than 1 year are conserved between the standard NEE and the flat prior, since the offsets are applied homogeneously in space and time within TSA and within 1 year. FLAT assimilates the data from the four surface sites in TSA in addition to those used by both MACCv10.1 and INVSAm. Of note is that even if increments on the NEE annual budget of a given year from an inversion are weak, the changes in the corresponding annual anomaly from the in-

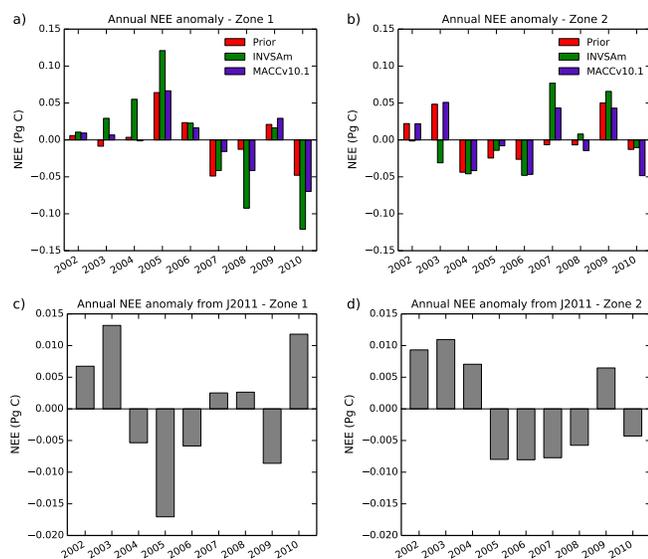


**Figure 9.** (a) Annual NEE anomaly compared to the mean of 2002–2010; estimates for the whole study region. (b) Annual NEE anomaly compared to the mean of 2002–2010; estimates for the area dominated by TBE forests.

version can be high because the inversion modifies the 2002–2010 average against which the anomaly is computed. Prior fluxes, MACCv10.1 and INVSAm display only small positive anomalies during the drought years (2005, 2010) compared to other years. FLAT displays a negative anomaly (i.e. a strong uptake) in 2010, but it indicates a larger positive anomaly in 2005 than that of other estimates. However, the strong NEE negative anomaly of 2009 in the prior fluxes, MACCv10.1 and INVSAm is also in FLAT, which suggests that this pattern is strongly driven by the atmospheric measurements and raises confidence in it.

As in Sect. 3.3.1, we isolated the results for the TBE forests area (Fig. 9b). In this case, prior fluxes and both MACCv10.1 and INVSAm estimates show diverging annual mean responses of forests to drought, with a positive anomaly in 2005 and a negative anomaly in 2010. For 2009, when climatic conditions were abnormally humid across South America, the inversion estimates consistently show a small positive anomaly, opposite to the response for the whole TSA region. The small anomalies in all inversions suggest a weak sensitivity of the NEE of TBE forests to interannual variations and that most of the IAV over the study area is not related to TBE forests.

Finally, we analyse the results in the two sub-regions shown in Fig. 8 in an attempt to identify potential differences in the regional responses. NEE estimates from the prior, INVSAm and MACCv10.1 show various responses of forests to drought in these zones. In Zone 1 (Fig. 10a) all these estimates present a positive anomaly in 2005 and a negative anomaly in 2010, while in Zone 2 (Fig. 10b) they yielded negative anomalies during both years. J2011 exhibit abnormal anomalies much smaller than these NEE estimates (Fig. 10c and d). This prevents us from gaining insights into the IAV from the comparison of J2011 to the other estimates.



**Figure 10.** Annual NEE anomaly compared to the 2002–2010 mean for Zone 1 (a, c) and Zone 2 (b, d) as defined in Fig. 8. Estimates from prior fluxes (red), INVSAm (green), MACCv10.1 (purple) and J2011 (grey).

However, the product of J2011 must be used cautiously, especially when evaluating IAV of NEE. J2011 relied on a limited number of EC stations across the Amazon basin, with short time series, to estimate MTE based on spatial gradients among the sites and then extrapolated to temporal gradients. This is valid assuming that spatial and temporal NEE patterns have the same sensitivity to climate, which may be incorrect (Piao et al., 2013). The example of the divergences of the results between MACCv10.1 and INVSAm in 2003 in Zone 2 illustrates, again, some weak ability to precisely constrain the fluxes in such a small area, which is quite distant from the measurement sites in TSA. Indeed, the analysis of the maps of increments from MACCv10.1 and INVSAm for the annual mean NEE in 2003 (not shown) demonstrates that the assimilation of data at SAN during this year shifts the northern border of the pattern of negative corrections in MACCv10.1 from north of Zone 2 to south of Zone 2. Since, on average over 2002–2010, both inversions apply positive increments in this zone (see Fig. 6), this leads to a clear negative annual anomaly in Zone 2 and for the year 2003 for INVSAm.

#### 4 Discussion and concluding remarks

Amazonian forests play a key role in the global carbon balance, but there are large uncertainties on the evolution of this terrestrial sink. Uncertainties stem from incomplete knowledge of the processes behind land–atmosphere CO<sub>2</sub> exchange in this region. Improving our understanding of the seasonal and interannual variations of Amazonian forests is thus a priority. In an attempt to gain insight into how these

temporal variations of CO<sub>2</sub> fluxes vary across Amazonia, we analysed global inversions and incorporated new measurements of atmospheric CO<sub>2</sub> mole fractions in TSA into one of these inversions. The analysis of the global inversions at such spatial scales, which are generally ignored in global inversion studies, is justified by the use of a variational inversion system solving for the fluxes at  $\sim 3^\circ$  and 8-day resolution. We showed that the two inversions applied large corrections to the estimates of NEE from a vegetation model that they used as prior information. The inverted NEE was strongly controlled by the assimilation of CO<sub>2</sub> measurements both outside and within the TSA region, and this control was characterized by zonal patterns of alternate positive and negative corrections, which we call “zonal dipole”, in addition to more local patterns in the vicinity of the sites that were assimilated.

Despite an overall improvement by the inversion of the seasonal variations of the simulated CO<sub>2</sub> mole fractions when compared to the measurements in TSA, several issues arose when analysing the seasonal cycles of NEE from the inversion. The seasonality of the mean NEE over the whole TSA region remained basically unchanged between the inversion estimates (Fig. 7a). The prior and inversion estimates of this mean seasonal cycle of NEE at the TSA scale are not in line with J2011 and disagree with the intuitive assumption that the seasonal cycle should be correlated with rainfall and solar radiation, especially in the tropical forest area. Furthermore, they do not exhibit a clear seasonal pattern over TBE forests at basin scale or within the analysed sub-regions. J2011 display a clear homogeneous seasonal cycle all the TSA region, which does not give confidence in its ability to distinguish regional heterogeneity. The proximity of Zone 1 to the stations in north-eastern Amazonia (SAN and GUY) (Fig. 8) suggests better confidence in the flux corrections applied by INVSAm to the prior fluxes in that zone than elsewhere in TSA region.

The reliability in the seasonal patterns of the inverted fluxes is thus not high, which seems to confirm that the zonal dipoles of increments from the inversion are artificial patterns that balance the overall correction in the Southern Hemisphere and are not necessarily consistent with the actual NEE in the TSA region. This is directly connected to the lack of CO<sub>2</sub> measurements in the TSA region both in space and time. The limited overlap among the TSA observations is a critical issue since measurements are often only available at a single site at once and, consequently, temporary model errors at this site can get far more weight in the inversion than if it had been balanced by information from other sites. Furthermore, the lack of confidence in the INVSAm results in Zone 1, which is relatively close to the GUY and SAN, suggests a low reliability in the statistics of the uncertainty in the prior NEE (in the inversion configuration), on which the extrapolation of the information from the vicinity of these sites to the whole north-east of the TSA region relies. This further supports the choice of avoiding computing posterior uncertainties in the inverted NEE as discussed in Sect. 2.1.

Such considerations also weaken the analysis of the IAV based on the inversion while J2011 do not provide a reliable IAV of the NEE in TSA, which could have supported such an analysis. However, some patterns of the IAV in the NEE seem consistent among the different inversion estimates when the atmospheric measurements have a strong control on it: across the TSA region the estimates from the prior fluxes, MACCv10.1, INVSAM and FLAT indicate small positive flux annual anomalies (CO<sub>2</sub> release) during the drought in 2005 and a strong negative (CO<sub>2</sub> sink) anomaly in 2009, presumably related to lower temperatures and more humid conditions in 2009. However, in 2010 there is a divergence of the results between the FLAT estimate and the others.

In the TBE forests, the highest source anomaly in 2005 seen in the prior fluxes, MACCv10.1 and INVSAM may be related to reduced photosynthesis during the drought, as found by Gatti et al. (2014), and/or tree mortality caused by the squall event of January 2005 (Negrón-Juárez et al., 2010). However, in 2010 these results indicate a small sink anomaly. This anomaly seems inconsistent with the hypothesis of a higher negative impact of the drought in 2010, which was more intense in terms of water stress and more geographically extensive (Lewis et al., 2011). However, it seems consistent with the recent results of Gatti et al. (2014), who found that the Amazon basin was carbon neutral during that year.

Even though some seasonal or interannual patterns from the inversion look realistic, our study mainly reveals some critical issues that hamper the ability to derive an accurate estimation of the temporal variability of NEE and of its spatial heterogeneity across Amazonian forests. A denser monitoring network across the basin with continuous time series, as initiated by Gatti et al. (2014), is needed to well constrain the fluxes in the region. In addition, the simulation of atmospheric transport may need to be handled with models that are better adapted to the local meteorological conditions. Regional transport models with higher spatial and temporal resolution and improved parameterisations of key atmospheric processes for the region (e.g. deep convection, Parazoo et al., 2008) have been developed (Moreira et al., 2013). The combination of a denser observation network and state-of-the-art regional modelling tools would overcome some of the critical limitations encountered here for the study of the temporal variability of biosphere CO<sub>2</sub> fluxes in Amazonia. Such regional inversion will require reliable regional configurations of the input error statistics, which could rely on extensions of the flux eddy-covariance measurement networks in Amazonia. Finally, adaptive strategies for the representation of the observations in the model simulations as a function of the sites and of the meteorological conditions (Law et al., 2010) could help loosen the selection of the data for the assimilation.

**The Supplement related to this article is available online at doi:10.5194/acp-15-8423-2015-supplement.**

*Acknowledgements.* We would like to thank Martin Jung (Max Planck Institute for Biogeochemistry) for the access to the upscaled NEE data. Data recorded at the GUY site were obtained in the framework of the GUYAFLUX project funded by the French Ministry of Research, INRA, and the CNES in the framework of the PO Feder Région Guyane. The GUYAFLUX project also received support from an Investissement d'Avenir grants of the French ANR (CEBA: ANR-10-LABX-0025). This study was co-funded by the European Commission under the EU Seventh Research Framework Programme (grant agreement no. 283080, Geocarbon project) and ARIA Technologies. G. Broquet acknowledges funding and support from the Chaire industrielle BridGES, a joint research program between Thales Alenia Space, Veolia, CEA, UVSQ and CNRS.

Edited by: R. Engelen

## References

- Araújo, A. C., Nobre, A. D., Kruijt, B., Elbers, J. A., Dallarosa, R., Stefani, P., von Randow, C., Manzi, A. O., Culf, A. D., Gash, J. H. C., Valentini, R., and Kabat, P.: Comparative measurements of carbon dioxide fluxes from two nearby towers in a central Amazonian rainforest: the Manaus LBA site, *J. Geophys. Res.*, 107, 8090, doi:10.1029/2001JD000676, 2002.
- Baker, I. T., Prihodko, L., Denning, A. S., Goulden, M., Miller, S., and da Rocha, H. R.: Seasonal drought stress in the Amazon: reconciling models and observations, *J. Geophys. Res.*, 113, G00B01, doi:10.1029/2007JG000644, 2008.
- Berrisford, P., Dee, D., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., and Uppala, S.: The ERA-Interim archive, Tech. rep., European Centre for Medium Range Weather Forecasts, Reading, 2009.
- Bonal, D., Bosc, A., Ponton, S., Goret, J.-Y., Burban, B., Gross, P., Bonnefond, J.-M., Elbers, J., Longdoz, B., Epron, D., Guehl, J.-M., and Granier, A.: Impact of severe dry season on net ecosystem exchange in the neotropical rainforest of French Guiana, *Glob. Change Biol.*, 14, 1917–1933, doi:10.1111/j.1365-2486.2008.01610.x, 2008.
- Butler, M. P., Davis, K. J., Denning, A. S., and Kawa, S. R.: Using continental observations in global atmospheric inversions of CO<sub>2</sub>: North American carbon sources and sinks, *Tellus B*, 62, 550–572, doi:10.1111/j.1600-0889.2010.00501.x, 2010.
- Chevallier, F. and O'Dell, C. W.: Error statistics of Bayesian CO<sub>2</sub> flux inversion schemes as seen from GOSAT, *Geophys. Res. Lett.*, 40, 1252–1256, doi:10.1002/grl.50228, 2013.
- Chevallier, F., Viovy, N., Reichstein, M., and Ciais, P.: On the assignment of prior errors in Bayesian inversions of CO<sub>2</sub> surface fluxes, *Geophys. Res. Lett.*, 33, L13802, doi:10.1029/2006GL026496, 2006.
- Chevallier, F., Bréon, F.-M., and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of

- CO<sub>2</sub> sources and sinks: Theoretical study in a variational data assimilation framework, *J. Geophys. Res.*, 112, D09307, doi:10.1029/2006JD007375, 2007.
- Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Fröhlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morguí, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO<sub>2</sub> surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements, *J. Geophys. Res.*, 115, D21307, doi:10.1029/2010JD013887, 2010.
- Chevallier, F., Wang, T., Ciais, P., Maignan, F., Bocquet, M., Altaf Arain, M., Cescatti, A., Chen, J., Dolman, A. J., Law, B. E., Margolis, H. A., Montagnani, L., and Moors, E. J.: What eddy-covariance measurements tell us about prior land flux errors in CO<sub>2</sub>-flux inversion schemes, *Global Biogeochem. Cy.*, 26, GB1021, doi:10.1029/2010GB003974, 2012.
- Ciais, P., Rayner, P., Chevallier, F., Bousquet, P., Logan, M., Peylin, P., and Ramonet, M.: Atmospheric inversions for estimating CO<sub>2</sub> fluxes: methods and perspectives, *Climatic Change*, 103, 69–92, 2010.
- Enting, I. G., Trudinger, C. M., and Francey, R. J.: A synthesis inversion of the concentration and  $\delta^{13}\text{C}$  of atmospheric CO<sub>2</sub>, *Tellus B*, 47, 35–52, 1995.
- Gatti, L. V., Miller, J. B., D'Amelio, M. T. S., Martinewski, A., Basso, L. S., Gloor, M. E., Wofsy, S., and Tans, P.: Vertical profiles of CO<sub>2</sub> above eastern Amazonia suggest a net carbon flux to the atmosphere and balanced biosphere between 2000 and 2009, *Tellus B*, 62, 581–594, doi:10.1111/j.1600-0889.2010.00484.x, 2010.
- Gatti, L. V., Gloor, M., Miller, J. B., Doughty, C. E., Malhi, Y., Domingues, L. G., Basso, L. S., Martinewski, A., Correia, C. S. C., Borges, V. F., Freitas, S., Braz, R., Anderson, L. O., Rocha, H., Grace, J., Phillips, O. L., and Lloyd, J.: Drought sensitivity of Amazonian carbon balance revealed by atmospheric measurements, *Nature*, 506, 76–80, doi:10.1038/nature12957, 2014.
- Gloor, M., Gatti, L., Brienen, R., Feldpausch, T. R., Phillips, O. L., Miller, J., Ometto, J. P., Rocha, H., Baker, T., de Jong, B., Houghton, R. A., Malhi, Y., Aragão, L. E. O. C., Guyot, J.-L., Zhao, K., Jackson, R., Peylin, P., Sitch, S., Poulter, B., Lomas, M., Zaehle, S., Huntingford, C., Levy, P., and Lloyd, J.: The carbon balance of South America: a review of the status, decadal trends and main determinants, *Biogeosciences*, 9, 5407–5430, doi:10.5194/bg-9-5407-2012, 2012.
- Grace, J., Malhi, Y., Lloyd, J., McIntyre, J., Miranda, A. C., Meir, P., and Miranda, H. S.: The use of eddy covariance to infer the net carbon dioxide uptake of Brazilian rain forest, *Glob. Change Biol.*, 2, 209–217, doi:10.1111/j.1365-2486.1996.tb00073.x, 1996.
- Hourdin, F., Musat, I., Bony, S., Braconnot, P., Codron, F., Dufresne, J.-L., Fairhead, L., Filiberti, M.-A., Friedlingstein, P., Grandpeix, J.-Y., Krinner, G., LeVan, P., Li, Z.-X., and Lott, F.: The LMDZ4 general circulation model: climate performance and sensitivity to parametrized physics with emphasis on tropical convection, *Clim. Dynam.*, 27, 787–813, doi:10.1007/s00382-006-0158-0, 2006.
- Huete, A. R., Didan, K., Shimabukuro, Y. E., Ratana, P., Saleska, S. R., Hutrya, L. R., Yang, W., Nemani, R. R., and Myneni, R.: Amazon rainforests green-up with sunlight in dry season, *Geophys. Res. Lett.*, 33, L06405, doi:10.1029/2005GL025583, 2006.
- INPE: Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite (PRODES), available at: <http://www.obt.inpe.br/prodes/index.php> (last access: 4 May 2014), 2011.
- Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.*, 116, G00J07, doi:10.1029/2010JG001566, 2011.
- Kirchhoff, V. W. J. H., Aires, C. B., and Alvala, P. C.: An experiment to determine atmospheric CO concentrations of tropical South Atlantic air samples, *Q. J. Roy. Meteorol. Soc.*, 129, 1891–1902, doi:10.1093/qj/129.633.1891, 2003.
- Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cy.*, 19, GB1015, doi:10.1029/2003GB002199, 2005.
- Law, R. M., Steele, L. P., Krummel, P. B., and Zahorowski, W.: Synoptic variations in atmospheric CO<sub>2</sub> at Cape Grim: a model intercomparison, *Tellus B*, 62, 810–820, doi:10.1111/j.1600-0889.2010.00470.x, 2010.
- Lewis, S. L., Brando, P. M., Phillips, O. L., van der Heijden, G. M. F., and Nepstad, D.: The 2010 Amazon drought, *Science*, 331, p. 554, doi:10.1126/science.1200807, 2011.
- Lloyd, J., Kolle, O., Fritsch, H., de Freitas, S. R., Silva Dias, M. A. F., Artaxo, P., Nobre, A. D., de Araújo, A. C., Kruijt, B., Sogacheva, L., Fisch, G., Thielmann, A., Kuhn, U., and Andreae, M. O.: An airborne regional carbon balance for Central Amazonia, *Biogeosciences*, 4, 759–768, doi:10.5194/bg-4-759-2007, 2007.
- Maignan, F., Bréon, F.-M., Chevallier, F., Viovy, N., Ciais, P., Garrec, C., Trules, J., and Mancip, M.: Evaluation of a Global Vegetation Model using time series of satellite vegetation indices, *Geosci. Model Dev.*, 4, 1103–1114, doi:10.5194/gmd-4-1103-2011, 2011.
- Malhi, Y., Roberts, J. T., Betts, R. A., Killeen, T. J., Li, W., and Nobre, C. A.: Climate change, deforestation, and the fate of the Amazon, *Science*, 319, 169–172, 2008.
- Malhi, Y., Aragão, L. E. O. C., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., Sitch, S., McSweeney, C., and Meir, P.: Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest, *P. Natl. Acad. Sci. USA*, 106, 20610–20615, doi:10.1073/pnas.0804619106, 2009.
- Marengo, J. A., Ronchail, J., Baez, J., and Alves, L.: State of the climate in 2009, in: *The Climate of Tropical South America East of the Andes*, vol. 91, American Meteorological Society, Boston, MA, USA, S148–S150, 2010.

- Moreira, D. S., Freitas, S. R., Bonatti, J. P., Mercado, L. M., Rosário, N. M. É., Longo, K. M., Miller, J. B., Gloor, M., and Gatti, L. V.: Coupling between the JULES land-surface scheme and the CCATT-BRAMS atmospheric chemistry model (JULES-CCATT-BRAMS1.0): applications to numerical weather forecasting and the CO<sub>2</sub> budget in South America, *Geosci. Model Dev.*, 6, 1243–1259, doi:10.5194/gmd-6-1243-2013, 2013.
- Morton, D. C., Nagol, J., Carabjal, C. C., Rosette, J., Palace, M., Cook, B. D., Vermote, E. F., Harding, D. J., and North, P. R. J.: Amazon forests maintain consistent canopy structure and greenness during the dry season, *Nature*, 506, 7487, doi:10.1038/nature13006, 2014.
- Negrón-Juárez, R. I., Chambers, J. Q., Guimaraes, G., Zeng, H., Raupp, C. F. M., Marra, D. M., Ribeiro, G. H. P. M., Saatchi, S. S., Nelson, B. W., and Higuchi, N.: Widespread Amazon forest tree mortality from a single cross-basin squall line event, *Geophys. Res. Lett.*, 37, L16701, doi:10.1029/2010GL043733, 2010.
- Olivier, J. and Bertoldo, J.: Global emissions sources and sinks, in: *The Climate System*, A. A. Balkema Publishers/Swets & Zeitlinger Publishers, Lisse, 33–78, 2001.
- Ometto, J. P. H. B., Nobre, A. D., Rocha, H. R., Artaxo, P., and Martinelli, L. A.: Amazonia and the modern carbon cycle: lessons learned, *Oecologia*, 143, 483–500, 2005.
- Parazoo, N. C., Denning, A. S., Kawa, S. R., Corbin, K. D., Lokupitiya, R. S., and Baker, I. T.: Mechanisms for synoptic variations of atmospheric CO<sub>2</sub> in North America, South America and Europe, *Atmos. Chem. Phys.*, 8, 7239–7254, doi:10.5194/acp-8-7239-2008, 2008.
- Parazoo, N. C., Bowman, K., Frankenberg, C., Lee, J.-E., Fisher, J. B., Worden, J., Jones, D. B. A., Berry, J., Collatz, G. J., Baker, I. T., Jung, M., Liu, J., Osterman, G., O'Dell, C., Sparks, A., Butz, A., Guerlet, S., Yoshida, Y., Chen, H., and Gerbig, C.: Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of XCO<sub>2</sub> and chlorophyll fluorescence from GOSAT, *Geophys. Res. Lett.*, 40, 2829–2833, doi:10.1002/grl.50452, 2013.
- Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO<sub>2</sub> data, *J. Geophys. Res.*, 107, 4385, doi:10.1029/2001JD000857, 2002.
- 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<sub>2</sub> inversions, *Biogeosciences*, 10, 6699–6720, doi:10.5194/bg-10-6699-2013, 2013.
- Phillips, O. L., Aragão, L. E. O. C., Lewis, S. L., Fisher, J. B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C. A., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T. R., Bánki, O., Blanc, L., Bonal, D., Brando, P., Chave, J., de Oliveira, A. C. A., Cardozo, N. D., Czimczik, C. I., Feldpausch, T. R., Freitas, M. A., Gloor, E., Higuchi, N., Jiménez, E., Lloyd, G., Meir, P., Mendoza, C., Morel, A., Neill, D. A., Nepstad, D., Patiño, S., Peñuela, M. C., Prieto, A., Ramírez, F., Schwarz, M., Silva, J., Silveira, M., Thomas, A. S., Steege, H. T., Stropp, J., Vásquez, R., Zelazowski, P., Dávila, E. A., Andelman, S., Andrade, A., Chao, K.-J., Erwin, T., Di Fiore, A. C. E. H., Keeling, H., Killeen, T. J., Laurance, W. F., Cruz, A. P., Pitman, N. C. A., Vargas, P. N., Ramírez-Angulo, H., Rudas, A., Salamão, R., Silva, N., Terborgh, J., and Torres-Lezama, A.: Drought sensitivity of the Amazon rainforest, *Science*, 323, 1344–1347, 2009.
- Piao, S., Sitch, S., Ciais, P., Friedlingstein, P., Peylin, P., Wang, X., Ahlström, A., Anav, A., Canadell, J. G., Cong, N., Huntingford, C., Jung, M., Levis, S., Levy, P. E., Li, J., Lin, X., Lomas, M. R., Lu, M., Luo, Y., Ma, Y., Myneni, R. B., Poulter, B., Sun, Z., Wang, T., Viovy, N., Zaehle, S., and Zeng, N.: Evaluation of terrestrial carbon cycle models for their response to climate variability and to CO<sub>2</sub> trends, *Glob. Chang. Biol.*, 19, 2117–2132, doi:10.1111/gcb.12187, 2013.
- Poulter, B., Heyder, U., and Cramer, W.: Modeling the sensitivity of the seasonal cycle of GPP to dynamic LAI and soil depths in tropical rainforests, *Ecosystems*, 12, 517–533, doi:10.1007/s10021-009-9238-4, 2009.
- Randerson, J. T., van der Werf, G. R., Giglio, L., Collatz, G. J., and Kasibhatla, P. S.: Global Fire Emissions Database, Version 2.1. Data set from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, doi:10.3334/ORNLDAAC/849, available at: <http://daac.ornl.gov/> (last access: 9 September 2013), 2007.
- Saatchi, S., Asefi-Najafabady, S., Malhi, Y., Aragão, L. E. O. C., Anderson, L. O., Myneni, R. B., and Nemani, R.: Persistent effects of a severe drought on Amazonian forest canopy, *P. Natl. Acad. Sci. USA*, 110, 565–570, 2012.
- Saleska, S. R., Miller, S. D., Matross, D. M., Goulden, M. L., Wofsy, S. C., da Rocha, H. R., de Camargo, P. B., Crill, P., Daube, B. C., de Freitas, H. C., Huttyra, L., Keller, M., Kirchhoff, V., Menton, M., Munger, J. W., Pyle, E. H., Rice, A. H., and Silva, H.: Carbon in Amazon forests: unexpected seasonal fluxes and disturbance-induced losses, *Science*, 302, 1554–1557, 2003.
- Saleska, S. R., Didan, K., Huete, A. R., and da Rocha, H. R.: Amazon forests green-up during 2005 drought, *Science*, 318, 612–612, 2007.
- Samanta, A., Ganguly, S., Hashimoto, H., Devadiga, S., Vermote, E., Knyazikhin, Y., Nemani, R. R., and Myneni, R. B.: Amazon forests did not green-up during the 2005 drought, *Geophys. Res. Lett.*, 37, L05401, doi:10.1029/2009GL042154, 2010.
- Samanta, A., Ganguly, S., Vermote, E., Nemani, R. R., and Myneni, R. B.: Interpretation of variations in MODIS-measured greenness levels of Amazon forests during 2000 to, *Environ. Res. Lett.*, 7, 024018, doi:10.1088/1748-9326/7/2/024018, 2012.
- Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C., Delille, B., Bates, N., and de Baar, H. J.: Climatological mean and decadal change in surface ocean pCO<sub>2</sub>, and net sea-air CO<sub>2</sub> flux over the global oceans, *Deep-Sea Res. Pt. II*, 56, 554–577, 2009.
- Tarantola, A.: *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, 2005.

- Verbeeck, H., Peylin, P., Bacour, C., Bonal, D., Steppe, K., and Ciais, P.: Seasonal patterns of CO<sub>2</sub> fluxes in Amazon forests: fusion of eddy covariance data and the ORCHIDEE model, *J. Geophys. Res.*, 116, G02018, doi:10.1029/2010JG001544, 2011.
- Wang, W., Ciais, P., Nemani, R. R., Canadell, J. G., Piao, S., Sitch, S., White, M. A., Hashimoto, H., Milesi, C., and Myneni, R. B.: Variations in atmospheric CO<sub>2</sub> growth rates coupled with tropical temperature, *P. Natl. Acad. Sci. USA*, 110, 13061–13066, doi:10.1073/pnas.1219683110, 2013.
- Xu, L., Samanta, A., Costa, M. H., Ganguly, S., Nemani, R. R., and Myneni, R. B.: Widespread decline in greenness of Amazonian vegetation due to the 2010 drought, *Geophys. Res. Lett.*, 38, L07402, doi:10.1029/2011GL046824, 2011.



APPENDIX

**B**

---

## Supplementary figures

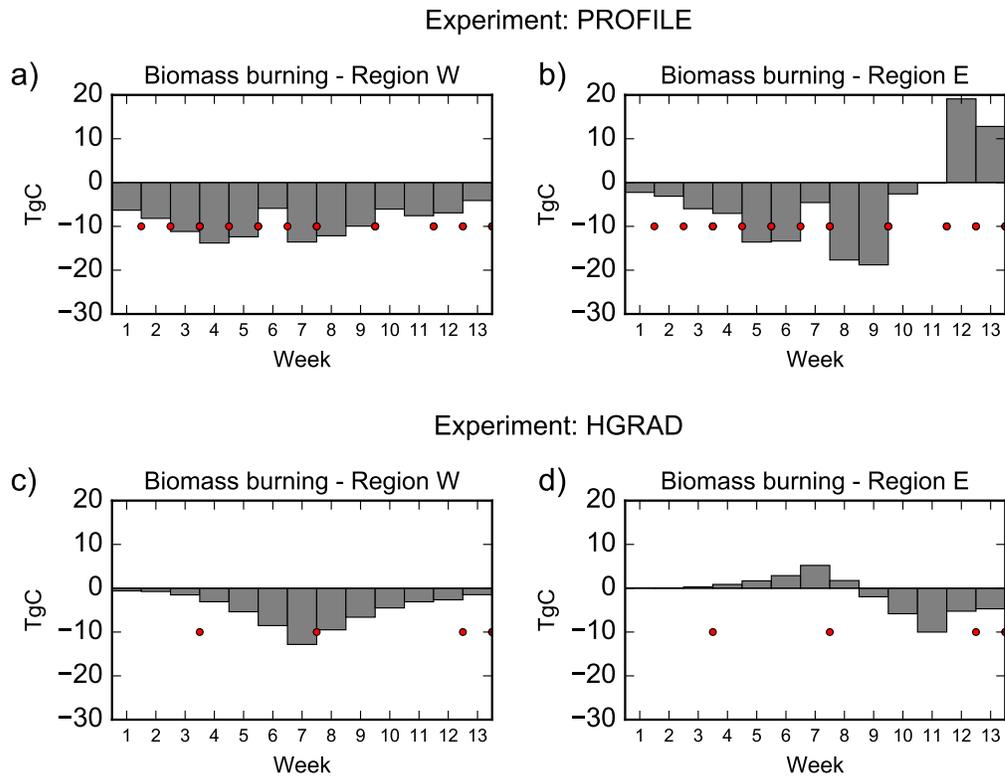


Figure B.1: Impact of the observations assimilated in the (a, b) PROFILE and (c, d) HGRAD experiments on the a posteriori weekly, sub-regional biomass burning CO<sub>2</sub> emissions, for the July – September (dry period) for INV-ECMWF. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observations were available.

Time series of flux corrections to NEE for every sub-region from INV-BRAMS  
Experiment: PROFILE

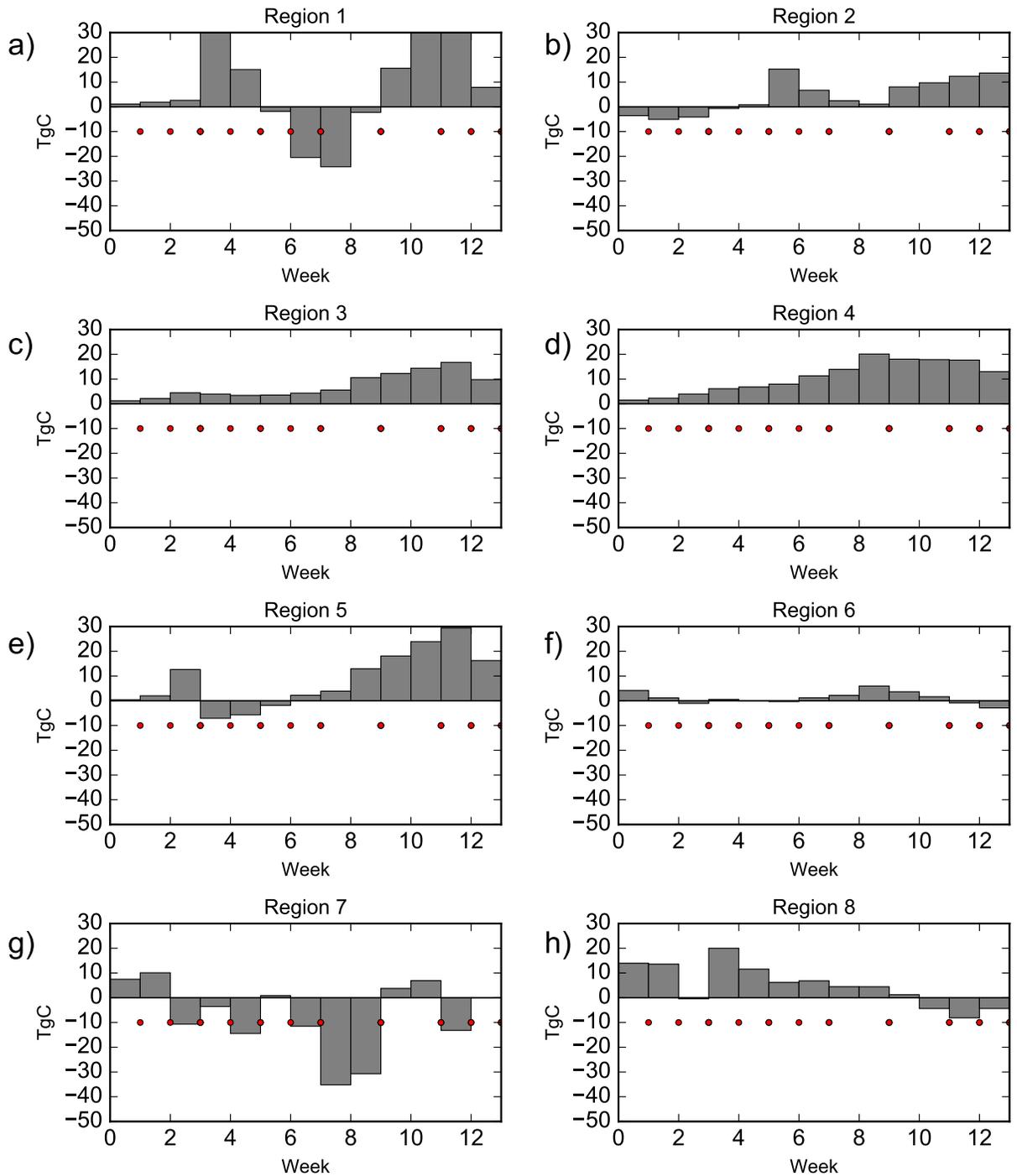


Figure B.2: Impact of the observations assimilated in the PROFILE experiment on the a posteriori weekly, sub-regional NEE, for the July – September (dry period) for INV-ECMWF. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observations were available.

Time series of flux corrections to NEE for every sub-region from INV-ECMWF  
Experiment: HGRAD

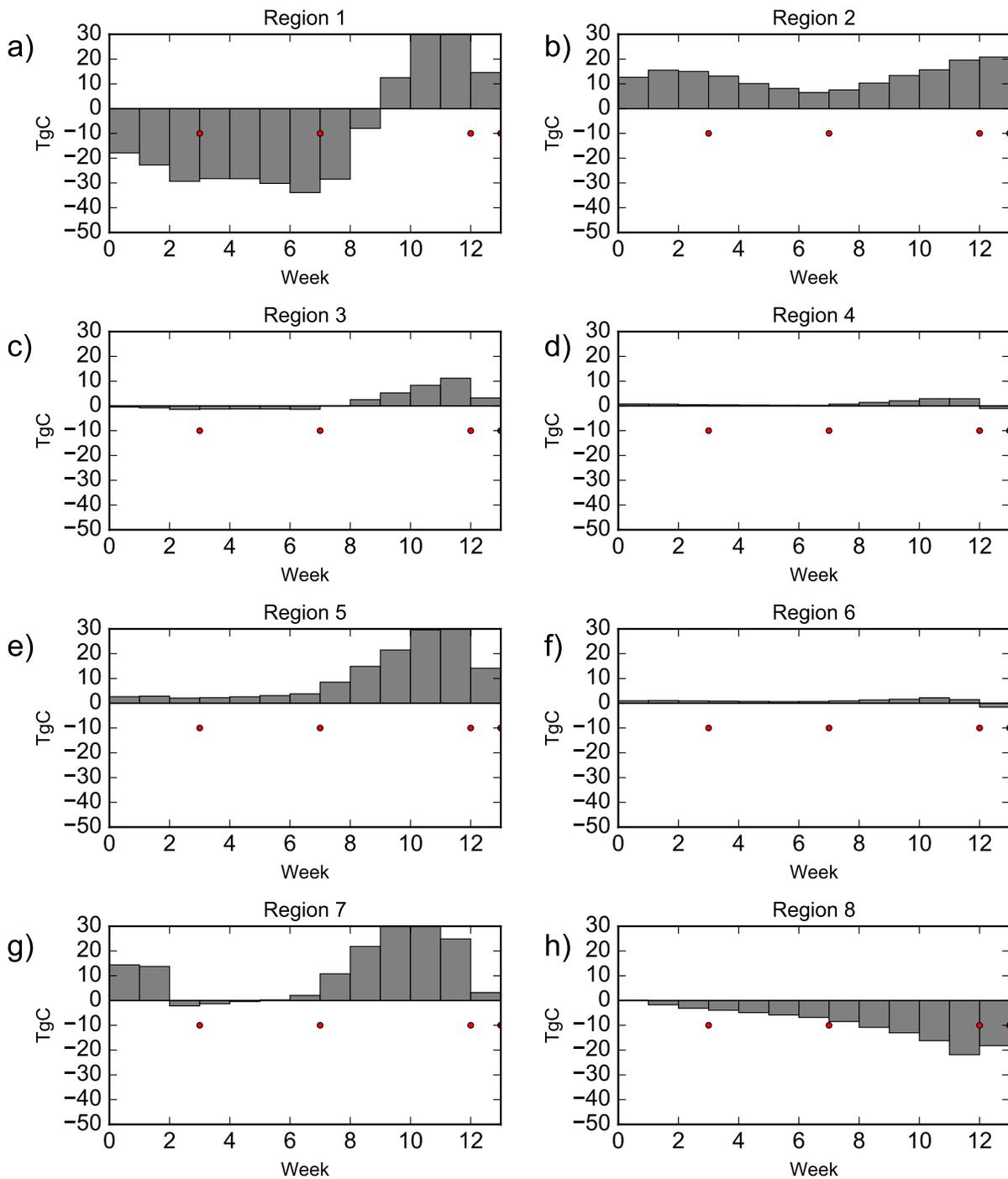


Figure B.3: Impact of the observations assimilated in the HGRAD experiment on the a posteriori weekly, sub-regional NEE, for the July – September (dry period) for INV-ECMWF. **Gray bars** are the flux correction to the weekly budget. **Red dots** indicate when observations were available.



---

## Bibliography

- Caroline B. Alden, John B. Miller, Luciana V. Gatti, Manuel M. Gloor, Kaiyu Guan, Anna M. Michalak, Ingrid T. van der Laan-Luijkx, Danielle Touma, Arlyn Andrews, Luana S. Basso, Caio S. C. Correia, Lucas G. Domingues, Joanna Joiner, Maarten C. Krol, Alexei I. Lyapunov, Wouter Peters, Yoichi P. Shiga, Kirk Thoning, Ivar R. van der Velde, Thijs T. van Leeuwen, Vineet Yadav, and Noah S. Diffenbaugh. Regional atmospheric CO<sub>2</sub> inversion reveals seasonal and geographic differences in Amazon net biome exchange. *Global Change Biology*, pages n/a–n/a, 2016. ISSN 1365-2486. doi: 10.1111/gcb.13305. URL <http://dx.doi.org/10.1111/gcb.13305>.
- Jeffrey S. Amthor. Terrestrial higher-plant response to increasing atmospheric [CO<sub>2</sub>] in relation to the global carbon cycle. *Global Change Biology*, 1(4):243–274, 1995. ISSN 1365-2486. doi: 10.1111/j.1365-2486.1995.tb00025.x. URL <http://dx.doi.org/10.1111/j.1365-2486.1995.tb00025.x>.
- M. O. Andreae, P. Artaxo, V. Beck, M. Bela, S. Freitas, C. Gerbig, K. Longo, J. W. Munger, K. T. Wiedemann, and S. C. Wofsy. Carbon monoxide and related trace gases and aerosols over the amazon basin during the wet and dry seasons. *Atmospheric Chemistry and Physics*, 12(13):6041–6065, 2012. doi: 10.5194/acp-12-6041-2012. URL <http://www.atmos-chem-phys.net/12/6041/2012/>.
- R. J. Andres, T. A. Boden, F.-M. Bréon, P. Ciais, S. Davis, D. Erickson, J. S. Gregg, A. Jacobson, G. Marland, J. Miller, T. Oda, J. G. J. Olivier, M. R. Raupach, P. Rayner, and K. Treanton. A synthesis of carbon dioxide emissions from fossil-fuel combustion. *Biogeosciences*, 9(5):1845–1871, 2012. doi: 10.5194/bg-9-1845-2012. URL <http://www.biogeosciences.net/9/1845/2012/>.
- Robert J. Andres, Gregg Marland, Inez Fung, and Elaine Matthews. A 1° × 1° distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. *Global Biogeochemical Cycles*, 10(3):419–429, 1996. ISSN 1944-9224. doi: 10.1029/96GB01523. URL <http://dx.doi.org/10.1029/96GB01523>.
- L. E. O. C. Aragão, Y. Malhi, D. B. Metcalfe, J. E. Silva-Espejo, E. Jiménez, D. Navarrete, S. Almeida, A. C. L. Costa, N. Salinas, O. L. Phillips, L. O. Anderson, E. Alvarez,

- T. R. Baker, P. H. Goncalvez, J. Huamán-Ovalle, M. Mamani-Solórzano, P. Meir, A. Monteagudo, S. Patiño, M. C. Peñuela, A. Prieto, C. A. Quesada, A. Rozas-Dávila, A. Rudas, J. A. Silva Jr., and R. Vásquez. Above- and below-ground net primary productivity across ten amazonian forests on contrasting soils. *Biogeosciences*, 6(12):2759–2778, 2009. doi: 10.5194/bg-6-2759-2009. URL <http://www.biogeosciences.net/6/2759/2009/>.
- D. C. E. Bakker, B. Pfeil, K. Smith, S. Hankin, A. Olsen, S. R. Alin, C. Cosca, S. Harasawa, A. Kozyr, Y. Nojiri, K. M. O'Brien, U. Schuster, M. Telszewski, B. Tilbrook, C. Wada, J. Akl, L. Barbero, N. R. Bates, J. Boutin, Y. Bozec, W.-J. Cai, R. D. Castle, F. P. Chavez, L. Chen, M. Chierici, K. Currie, H. J. W. de Baar, W. Evans, R. A. Feely, A. Fransson, Z. Gao, B. Hales, N. J. Hardman-Mountford, M. Hoppema, W.-J. Huang, C. W. Hunt, B. Huss, T. Ichikawa, T. Johannessen, E. M. Jones, S. D. Jones, S. Jutterström, V. Kitidis, A. Körtzinger, P. Landschützer, S. K. Lauvset, N. Lefèvre, A. B. Manke, J. T. Mathis, L. Merlivat, N. Metzl, A. Murata, T. Newberger, A. M. Omar, T. Ono, G.-H. Park, K. Paterson, D. Pierrot, A. F. Ríos, C. L. Sabine, S. Saito, J. Salisbury, V. V. S. S. Sarma, R. Schlitzer, R. Sieger, I. Skjelvan, T. Steinhoff, K. F. Sullivan, H. Sun, A. J. Sutton, T. Suzuki, C. Sweeney, T. Takahashi, J. Tjiputra, N. Tsurushima, S. M. A. C. van Heuven, D. Vandemark, P. Vlahos, D. W. R. Wallace, R. Wanninkhof, and A. J. Watson. An update to the surface ocean  $\text{CO}_2$  atlas (socat version 2). *Earth System Science Data*, 6(1):69–90, 2014. doi: 10.5194/essd-6-69-2014. URL <http://www.earth-syst-sci-data.net/6/69/2014/>.
- N. H. Batjes. Documentation to isric-wise global data set of derived soil properties on a 1/2 deg by 1/2 deg grid (version 1.0). Working paper and preprint 96/05, ISRIC, Wageningen, 1996.
- M. M. Bela, K. M. Longo, S. R. Freitas, D. S. Moreira, V. Beck, S. C. Wofsy, C. Gerbig, K. Wiedemann, M. O. Andreae, and P. Artaxo. Ozone production and transport over the amazon basin during the dry-to-wet and wet-to-dry transition seasons. *Atmospheric Chemistry and Physics*, 15(2):757–782, 2015. doi: 10.5194/acp-15-757-2015. URL <http://www.atmos-chem-phys.net/15/757/2015/>.
- D. A. Belikov, S. Maksyutov, A. Ganshin, R. Zhuravlev, N. M. Deutscher, D. Wunch, D. G. Feist, I. Morino, R. J. Parker, K. Strong, Y. Yoshida, A. Bril, S. Oshchepkov, H. Boesch, M. K. Dubey, D. Griffith, W. Hewson, R. Kivi, J. Mendonca, J. Notholt, M. Schneider, R. Sussmann, V. A. Velazco, and S. Aoki. Study of the footprints of short-term variation in  $\text{XCO}_2$  observed by TCCON sites using NIES and FLEXPART atmospheric transport models. *Atmospheric Chemistry and Physics*, 17(1):143–157, 2017. doi: 10.5194/acp-17-143-2017. URL <http://www.atmos-chem-phys.net/17/143/2017/>.
- P. Berrisford, D. Dee, K. Fielding, M. Fuentes, P. Kallberg, S. Kobayashi, and S. Uppala. The era-interim archive. Technical report, European Centre for Medium Range Weather Forecasts, Reading, 2009.
- M. J. Best, M. Pryor, D. B. Clark, G. G. Rooney, R. L. H. Essery, C. B. Ménard, J. M. Edwards, M. A. Hendry, A. Porson, N. Gedney, L. M. Mercado, S. Sitch, E. Blyth, O. Boucher, P. M. Cox, C. S. B. Grimmond, and R. J. Harding. The joint uk land environment simulator (jules), model description – part 1: Energy and water fluxes. *Geoscientific Model Development*, 4(3):677–699, 2011. doi: 10.5194/gmd-4-677-2011. URL <http://www.geosci-model-dev.net/4/677/2011/>.

- T. A. Boden, G. Marland, and R. J. Andres. Global, regional, and national fossil-fuel CO<sub>2</sub> emissions. Technical report, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., USA, 2013.
- B. Bolin and C. D. Keeling. Large-scale atmospheric mixing as deduced from the seasonal and meridional variations of carbon dioxide. *Journal of Geophysical Research*, 68(13): 3899–3920, 1963. ISSN 2156-2202. doi: 10.1029/JZ068i013p03899. URL <http://dx.doi.org/10.1029/JZ068i013p03899>.
- Gordon B. Bonan. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, 320(5882):1444–1449, 2008. ISSN 0036-8075. doi: 10.1126/science.1155121. URL <http://science.sciencemag.org/content/320/5882/1444>.
- A. Boon, G. Broquet, D. J. Clifford, F. Chevallier, D. M. Butterfield, I. Pison, M. Ramonet, J.-D. Paris, and P. Ciais. Analysis of the potential of near-ground measurements of CO<sub>2</sub> and CH<sub>4</sub> in London, UK, for the monitoring of city-scale emissions using an atmospheric transport model. *Atmospheric Chemistry and Physics*, 16(11):6735–6756, 2016. doi: 10.5194/acp-16-6735-2016. URL <http://www.atmos-chem-phys.net/16/6735/2016/>.
- Philippe Bousquet, Philippe Peylin, Philippe Ciais, Corinne Le Quéré, Pierre Friedlingstein, and Pieter P. Tans. Regional Changes in Carbon Dioxide Fluxes of Land and Oceans Since 1980. *Science*, 290(5495):1342–1346, 2000. ISSN 0036-8075. doi: 10.1126/science.290.5495.1342. URL <http://science.sciencemag.org/content/290/5495/1342>.
- F. Bouttier and P. Courtier. Data assimilation concepts and methods, 2002.
- F. M. Bréon, G. Broquet, V. Puygrenier, F. Chevallier, I. Xueref-Remy, M. Ramonet, E. Dieudonné, M. Lopez, M. Schmidt, O. Perrussel, and P. Ciais. An attempt at estimating paris area CO<sub>2</sub> emissions from atmospheric concentration measurements. *Atmospheric Chemistry and Physics*, 15(4):1707–1724, 2015. doi: 10.5194/acp-15-1707-2015. URL <http://www.atmos-chem-phys.net/15/1707/2015/>.
- Grégoire Broquet, Frédéric Chevallier, Peter Rayner, Céline Aulagnier, Isabelle Pison, Michel Ramonet, Martina Schmidt, Alex T. Vermeulen, and Philippe Ciais. A european summertime CO<sub>2</sub> biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements. *J. Geophys. Res.*, 116(D23):p. D23303, December 2011. ISSN 2156-2202. doi: 10.1029/2011JD016202. URL <http://dx.doi.org/10.1029/2011JD016202>.
- M. Buchwitz, I. Khlystova, H. Bovensmann, and J. P. Burrows. Three years of global carbon monoxide from sciamachy: comparison with mopitt and first results related to the detection of enhanced CO over cities. *Atmospheric Chemistry and Physics*, 7(9):2399–2411, 2007. doi: 10.5194/acp-7-2399-2007. URL <http://www.atmos-chem-phys.net/7/2399/2007/>.
- Josep G. Canadell, Philippe Ciais, Kevin Gurney, Corinne Le Quéré, Shilong Piao, Michael R. Raupach, and Christopher L. Sabine. An international effort to quantify regional carbon fluxes. *Eos, Transactions American Geophysical Union*, 92(10):81–82, 2011. ISSN 2324-9250. doi: 10.1029/2011EO100001. URL <http://dx.doi.org/10.1029/2011EO100001>.

- Mingkui Cao and F. Ian Woodward. Dynamic responses of terrestrial ecosystem carbon cycling to global climate change. *Nature*, 393(6682):249–252, May 1998. ISSN 0028-0836. URL <http://dx.doi.org/10.1038/30460>.
- CEOS. CEOS Strategy for Carbon Observations from Space. *The Committee on Earth Observation Satellites (CEOS) Response to the Group on Earth Observations (GEO) Carbon Strategy*, 2014.
- F. Chevallier. On the statistical optimality of  $\text{CO}_2$  atmospheric inversions assimilating  $\text{CO}_2$  column retrievals. *Atmospheric Chemistry and Physics*, 15(19):11133–11145, 2015. doi: 10.5194/acp-15-11133-2015. URL <http://www.atmos-chem-phys.net/15/11133/2015/>.
- F. Chevallier, M. Fisher, P. Peylin, S. Serrar, P. Bousquet, F.-M. Bréon, A. Chédin, and P. Ciais. Inferring  $\text{CO}_2$  sources and sinks from satellite observations: Method and application to tovs data. *Journal of Geophysical Research: Atmospheres*, 110(D24):n/a–n/a, 2005. ISSN 2156-2202. doi: 10.1029/2005JD006390. URL <http://dx.doi.org/10.1029/2005JD006390>.
- F. Chevallier, P. Ciais, T. J. Conway, T. Aalto, B. E. Anderson, P. Bousquet, E. G. Brunke, L. Ciattaglia, Y. Esaki, M. Fröhlich, A. Gomez, A. J. Gomez-Pelaez, L. Haszpra, P. B. Krummel, R. L. Langenfelds, M. Leuenberger, T. Machida, F. Maignan, H. Matsueda, J. A. Morguí, H. Mukai, T. Nakazawa, P. Peylin, M. Ramonet, L. Rivier, Y. Sawa, M. Schmidt, L. P. Steele, S. A. Vay, A. T. Vermeulen, S. Wofsy, and D. Worthy.  $\text{CO}_2$  surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements. *J. Geophys. Res.*, 115(D21):D21307–, November 2010. ISSN 2156-2202. URL <http://dx.doi.org/10.1029/2010JD013887>.
- Frédéric Chevallier and Christopher W. O’Dell. Error statistics of bayesian  $\text{CO}_2$  flux inversion schemes as seen from gosat. *Geophysical Research Letters*, 40(6):1252–1256, 2013. ISSN 1944-8007. doi: 10.1002/grl.50228. URL <http://dx.doi.org/10.1002/grl.50228>.
- Frédéric Chevallier, Shamil Maksyutov, Philippe Bousquet, François-Marie Bréon, Ryu Saito, Yukio Yoshida, and Tatsuya Yokota. On the accuracy of the  $\text{CO}_2$  surface fluxes to be estimated from the gosat observations. *Geophysical Research Letters*, 36(19):n/a–n/a, 2009. ISSN 1944-8007. doi: 10.1029/2009GL040108. URL <http://dx.doi.org/10.1029/2009GL040108>.
- Frédéric Chevallier, Paul I. Palmer, Liang Feng, Hartmut Boesch, Christopher W. O’Dell, and Philippe Bousquet. Toward robust and consistent regional  $\text{CO}_2$  flux estimates from in situ and spaceborne measurements of atmospheric  $\text{CO}_2$ . *Geophysical Research Letters*, 41(3):1065–1070, 2014. ISSN 1944-8007. doi: 10.1002/2013GL058772. URL <http://dx.doi.org/10.1002/2013GL058772>. 2013GL058772.
- Wendy W. Chou, Steven C. Wofsy, Robert C. Harriss, John C. Lin, C. Gerbig, and Glenn W. Sachse. Net fluxes of  $\text{CO}_2$  in amazonia derived from aircraft observations. *Journal of Geophysical Research: Atmospheres*, 107(D22):ACH 4–1–ACH 4–15, 2002. ISSN 2156-2202. doi: 10.1029/2001JD001295. URL <http://dx.doi.org/10.1029/2001JD001295>.
- P. Ciais, A. V. Borges, G. Abril, M. Meybeck, G. Folberth, D. Hauglustaine, and I. A. Janssens. The impact of lateral carbon fluxes on the European carbon balance. *Biogeosciences*, 5(5):

- 1259–1271, 2008. doi: 10.5194/bg-5-1259-2008. URL <http://www.biogeosciences.net/5/1259/2008/>.
- P. Ciais, P. Rayner, F. Chevallier, P. Bousquet, M. Logan, P. Peylin, and M. Ramonet. Atmospheric inversions for estimating CO<sub>2</sub> fluxes: methods and perspectives. *Clim. Chang.*, 103(1-2):69–92, July 2010.
- D. B. Clark, L. M. Mercado, S. Sitch, C. D. Jones, N. Gedney, M. J. Best, M. Pryor, G. G. Rooney, R. L. H. Essery, E. Blyth, O. Boucher, R. J. Harding, C. Huntingford, and P. M. Cox. The joint uk land environment simulator (jules), model description – part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3):701–722, 2011. doi: 10.5194/gmd-4-701-2011. URL <http://www.geosci-model-dev.net/4/701/2011/>.
- C. Crevoisier, A. Chédin, H. Matsueda, T. Machida, R. Armante, and N. A. Scott. First year of upper tropospheric integrated content of co<sub>2</sub> from iasi hyperspectral infrared observations. *Atmospheric Chemistry and Physics*, 9(14):4797–4810, 2009. doi: 10.5194/acp-9-4797-2009. URL <http://www.atmos-chem-phys.net/9/4797/2009/>.
- D Crisp, R.M Atlas, F.-M Breon, L.R Brown, J.P Burrows, P Ciais, B.J Connor, S.C Doney, I.Y Fung, D.J Jacob, C.E Miller, D O’Brien, S Pawson, J.T Randerson, P Rayner, R.J Salawitch, S.P Sander, B Sen, G.L Stephens, P.P Tans, G.C Toon, P.O Wennberg, S.C Wofsy, Y.L Yung, Z Kuang, B Chudasama, G Sprague, B Weiss, R Pollock, D Kenyon, and S Schroll. The orbiting carbon observatory (oco) mission. *Advances in Space Research*, 34(4):700 – 709, 2004. ISSN 0273-1177. doi: <http://dx.doi.org/10.1016/j.asr.2003.08.062>. URL <http://www.sciencedirect.com/science/article/pii/S0273117704003539>. Trace Constituents in the Troposphere and Lower Stratosphere.
- F. Deng, D. B. A. Jones, D. K. Henze, N. Bousserez, K. W. Bowman, J. B. Fisher, R. Nassar, C. O’Dell, D. Wunch, P. O. Wennberg, E. A. Kort, S. C. Wofsy, T. Blumenstock, N. M. Deutscher, D. W. T. Griffith, F. Hase, P. Heikkinen, V. Sherlock, K. Strong, R. Sussmann, and T. Warneke. Inferring regional sources and sinks of atmospheric co<sub>2</sub> from gosat xco<sub>2</sub> data. *Atmospheric Chemistry and Physics*, 14(7):3703–3727, 2014. doi: 10.5194/acp-14-3703-2014. URL <http://www.atmos-chem-phys.net/14/3703/2014/>.
- G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3385–3396, 2005. ISSN 1477-870X. doi: 10.1256/qj.05.108. URL <http://dx.doi.org/10.1256/qj.05.108>.
- I. G. Enting. *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, 2002.
- EU-JRC/PBL. Emission Database for Global Atmospheric Research (EDGAR), Release Version 4.2 FastTrack 2010, 2013. URL <http://edgar.jrc.ec.europa.eu>. Available at.
- S. Fan, M. Gloor, J. Mahlman, S. Pacala, J. Sarmiento, T. Takahashi, and P. Tans. A Large Terrestrial Carbon Sink in North America Implied by Atmospheric and Oceanic Carbon Dioxide Data and Models. *Science*, 282(5388):442–446, 1998. ISSN 0036-8075. doi: 10.1126/science.282.5388.442. URL <http://science.sciencemag.org/content/282/5388/442>.

- Simone Fatichi, Sebastian Leuzinger, Athanasios Paschalis, J. Adam Langley, Alicia Donnellan Barraclough, and Mark J. Hovenden. Partitioning direct and indirect effects reveals the response of water-limited ecosystems to elevated CO<sub>2</sub>. *Proceedings of the National Academy of Sciences*, 113(45):12757–12762, 2016. doi: 10.1073/pnas.1605036113. URL <http://www.pnas.org/content/113/45/12757.abstract>.
- T. R. Feldpausch, J. Lloyd, S. L. Lewis, R. J. W. Brienen, M. Gloor, A. Monteagudo Mendoza, G. Lopez-Gonzalez, L. Banin, K. Abu Salim, K. Affum-Baffoe, M. Alexiades, S. Almeida, I. Amaral, A. Andrade, L. E. O. C. Aragão, A. Araujo Murakami, E. J. M. M. Arets, L. Arroyo, G. A. Aymard C., T. R. Baker, O. S. Bánki, N. J. Berry, N. Cardozo, J. Chave, J. A. Comiskey, E. Alvarez, A. de Oliveira, A. Di Fiore, G. Djagbletey, T. F. Domingues, T. L. Erwin, P. M. Fearnside, M. B. França, M. A. Freitas, N. Higuchi, E. Honorio C., Y. Iida, E. Jiménez, A. R. Kassim, T. J. Killeen, W. F. Laurance, J. C. Lovett, Y. Malhi, B. S. Marimon, B. H. Marimon-Junior, E. Lenza, A. R. Marshall, C. Mendoza, D. J. Metcalfe, E. T. A. Mitchard, D. A. Neill, B. W. Nelson, R. Nilus, E. M. Nogueira, A. Parada, K. S.-H. Peh, A. Pena Cruz, M. C. Peñuela, N. C. A. Pitman, A. Prieto, C. A. Quesada, F. Ramírez, H. Ramírez-Angulo, J. M. Reitsma, A. Rudas, G. Saiz, R. P. Salomão, M. Schwarz, N. Silva, J. E. Silva-Espejo, M. Silveira, B. Sonké, J. Stropp, H. E. Taedoumg, S. Tan, H. ter Steege, J. Terborgh, M. Torello-Raventos, G. M. F. van der Heijden, R. Vásquez, E. Vilanova, V. A. Vos, L. White, S. Willcock, H. Woell, and O. L. Phillips. Tree height integrated into pantropical forest biomass estimates. *Biogeosciences*, 9(8):3381–3403, 2012. doi: 10.5194/bg-9-3381-2012. URL <http://www.biogeosciences.net/9/3381/2012/>.
- L. Feng, P. I. Palmer, H. Bösch, R. J. Parker, A. J. Webb, C. S. C. Correia, N. M. Deutscher, L. G. Domingues, D. G. Feist, L. V. Gatti, E. Gloor, F. Hase, R. Kivi, Y. Liu, J. B. Miller, I. Morino, R. Sussmann, K. Strong, O. Uchino, J. Wang, and A. Zahn. Consistent regional fluxes of ch<sub>4</sub> and co<sub>2</sub> inferred from gosat proxy xch<sub>4</sub>:xco<sub>2</sub> retrievals, 2010–2014. *Atmospheric Chemistry and Physics*, 17(7):4781–4797, 2017. doi: 10.5194/acp-17-4781-2017. URL <http://www.atmos-chem-phys.net/17/4781/2017/>.
- E. J. Fiktau. Esboço de uma divisão ecológica da região amazônica. In *Proc. Symp. Biol. Trop. Amaz.* Florencia y Leticia, 1969.
- G. Fisch, J. Tota, T. L. A. Machado, F. M. A. Silva Dias, F. R. da F. Lyra, A. C. Nobre, J. A. Dolman, and C. J. H. Gash. The convective boundary layer over pasture and forest in amazonia. *Theoretical and Applied Climatology*, 78(1):47–59, 2004. ISSN 1434-4483. doi: 10.1007/s00704-004-0043-x. URL <http://dx.doi.org/10.1007/s00704-004-0043-x>.
- S. R. Freitas, K. M. Longo, M. A. F. Silva Dias, R. Chatfield, P. Silva Dias, P. Artaxo, M. O. Andreae, G. Grell, L. F. Rodrigues, A. Fazenda, and J. Panetta. The coupled aerosol and tracer transport model to the brazilian developments on the regional atmospheric modeling system (catt-brams) - part 1: Model description and evaluation. *Atmos. Chem. Phys.*, 9(8):2843–2861, April 2009. ISSN 1680-7324. URL <http://www.atmos-chem-phys.net/9/2843/2009/>.
- Mark A. Friedl, Damien Sulla-Menashe, Bin Tan, Annemarie Schneider, Navin Ramankutty, Adam Sibley, and Xiaoman Huang. {MODIS} collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment*, 114(1):168 – 182, 2010. ISSN 0034-4257. doi: <http://dx.doi.org/>

- 10.1016/j.rse.2009.08.016. URL <http://www.sciencedirect.com/science/article/pii/S0034425709002673>.
- René D. Garreaud, Mathias Vuille, Rosa Compagnucci, and José Marengo. Present-day south american climate. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 281(3–4):180 – 195, 2009. ISSN 0031-0182. doi: <http://dx.doi.org/10.1016/j.palaeo.2007.10.032>. URL <http://www.sciencedirect.com/science/article/pii/S0031018208005002>. Long-term multi-proxy climate reconstructions and dynamics in South America (LOTRED-SA): State of the art and perspectives.
- T. Gasser and P. Ciais. A theoretical framework for the net land-to-atmosphere  $\text{CO}_2$  flux and its implications in the definition of "emissions from land-use change". *Earth System Dynamics*, 4(1):171–186, 2013. doi: 10.5194/esd-4-171-2013. URL <http://www.earth-syst-dynam.net/4/171/2013/>.
- L. V. Gatti, J. B. Miller, M. T. S. D'Amelio, A. Martinewski, L. S. Basso, M. E. Gloor, S. Wofsy, and P. Tans. Vertical profiles of  $\text{CO}_2$  above eastern Amazonia suggest a net carbon flux to the atmosphere and balanced biosphere between 2000 and 2009. *Tellus B*, 62(5): 581–594, 2010. ISSN 1600-0889. doi: 10.1111/j.1600-0889.2010.00484.x. URL <http://dx.doi.org/10.1111/j.1600-0889.2010.00484.x>.
- L. V. Gatti, M. Gloor, J. B. Miller, C. E. Doughty, Y. Malhi, L. G. Domingues, L. S. Basso, A. Martinewski, C. S. C. Correia, V. F. Borges, S. Freitas, R. Braz, L. O. Anderson, H. Rocha, J. Grace, O. L. Phillips, and J. Lloyd. Drought sensitivity of amazonian carbon balance revealed by atmospheric measurements. *Nature*, 506(7486):76–80, February 2014. ISSN 0028-0836. URL <http://dx.doi.org/10.1038/nature12957>.
- C. Geels, M. Gloor, P. Ciais, P. Bousquet, P. Peylin, A. T. Vermeulen, R. Dargaville, T. Aalto, J. Brandt, J. H. Christensen, L. M. Frohn, L. Haszpra, U. Karstens, C. Rödenbeck, M. Ramonet, G. Carboni, and R. Santaguida. Comparing atmospheric transport models for future regional inversions over europe &ndash; part 1: mapping the atmospheric  $\text{CO}_2$  signals. *Atmospheric Chemistry and Physics*, 7(13):3461–3479, 2007. doi: 10.5194/acp-7-3461-2007. URL <http://www.atmos-chem-phys.net/7/3461/2007/>.
- Dean B. Gesch, Kristine L. Verdin, and Susan K. Greenlee. New land surface digital elevation model covers the Earth. *Eos, Transactions American Geophysical Union*, 80(6): 69–70, 1999. ISSN 2324-9250. doi: 10.1029/99EO00050. URL <http://dx.doi.org/10.1029/99EO00050>.
- R. Gevaerd and S. R. Freitas. Estimativa operacional da umidade do solo para inicialização de modelos de previsão numérica da atmosfera. Parte i: Descrição da metodologia e validação. *Revista Brasileira de Meteorologia*, 21(3):1–15, 2006.
- Louis Giglio, James T. Randerson, and Guido R. van der Werf. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences*, 118(1):317–328, 2013. ISSN 2169-8961. doi: 10.1002/jgrg.20042. URL <http://dx.doi.org/10.1002/jgrg.20042>.
- GLOBALVIEW-CO2. Cooperative Global Atmospheric Data Integration Project, Multi-laboratory compilation of synchronized and gap-filled atmospheric carbon dioxide records for the period 1979-2012 (obspack\_co2\_1\_GLOBALVIEW-

- CO2\_2013\_1.0.4\_2013-12-23), update annually. NOAA Global Monitoring Division, Boulder, Colorado, U.S.A, 2013. doi: 10.3334/OBSPACK/1002.
- M. Gloor, L. Gatti, R. Brienen, T. R. Feldpausch, O. L. Phillips, J. Miller, J. P. Ometto, H. Rocha, T. Baker, B. de Jong, R. A. Houghton, Y. Malhi, L. E. O. C. Aragão, J.-L. Guyot, K. Zhao, R. Jackson, P. Peylin, S. Sitch, B. Poulter, M. Lomas, S. Zaehle, C. Huntingford, P. Levy, and J. Lloyd. The carbon balance of south america: a review of the status, decadal trends and main determinants. *Biogeosciences*, 9(12):5407–5430, 2012. doi: 10.5194/bg-9-5407-2012. URL <http://www.biogeosciences.net/9/5407/2012/>.
- Manuel Gloor, Song-Miao Fan, Stephen Pacala, and Jorge Sarmiento. Optimal sampling of the atmosphere for purpose of inverse modeling: A model study. *Global Biogeochemical Cycles*, 14(1):407–428, 2000. ISSN 1944-9224. doi: 10.1029/1999GB900052. URL <http://dx.doi.org/10.1029/1999GB900052>.
- KEVIN ROBERT Gurney, RACHEL M. Law, A. SCOTT Denning, PETER J. Rayner, DAVID Baker, PHILIPPE Bousquet, LORI Bruhwiler, YU-HAN Chen, PHILIPPE Ciais, SONGMIAO Fan, INEZ Y. Fung, MANUEL Gloor, MARTIN Heimann, KAZ Higuchi, JASMIN John, EVA Kowalczyk, Takashi Maki, SHAMIL Maksyutov, Philippe Peylin, MICHAEL Prather, BERNARD C. Pak, JORGE Sarmiento, SHOICHI Taguchi, TARO Takahashi, and CHIU-WAI Yuen. Transcom 3 co2 inversion intercomparison: 1. annual mean control results and sensitivity to transport and prior flux information. *Tellus B*, 55(2):555–579, 2003. ISSN 1600-0889. doi: 10.1034/j.1600-0889.2003.00049.x. URL <http://dx.doi.org/10.1034/j.1600-0889.2003.00049.x>.
- Valdir Herrmann and Saulo Ribeiro Freitas. ATMOSPHERIC CO<sub>2</sub> BUDGET OVER THE AMAZON BASIN:THE ROLE OF CONVECTIVE SYSTEMS. *Revista Brasileira de Meteorologia*, 26(4):529–540, 2011.
- J. R Holton. *An Introduction to Dynamic Meteorology*. Elsevier, fourth edition, 2004.
- P. B. Hooghiemstra, M. C. Krol, J. F. Meirink, P. Bergamaschi, G. R. van der Werf, P. C. Novelli, I. Aben, and T. Röckmann. Optimizing global co emission estimates using a four-dimensional variational data assimilation system and surface network observations. *Atmospheric Chemistry and Physics*, 11(10):4705–4723, 2011. doi: 10.5194/acp-11-4705-2011. URL <http://www.atmos-chem-phys.net/11/4705/2011/>.
- Alfredo R. Huete, Kamel Didan, Yosio E. Shimabukuro, Piyachat Ratana, Scott R. Saleska, Lucy R. Hutyrá, Wenze Yang, Ramakrishna R. Nemani, and Ranga Myneni. Amazon rainforests green-up with sunlight in dry season. *Geophys. Res. Lett.*, 33(6):L06405–, March 2006. ISSN 1944-8007. URL <http://dx.doi.org/10.1029/2005GL025583>.
- IPCC. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2013.
- Martin Jung, Markus Reichstein, Hank A. Margolis, Alessandro Cescatti, Andrew D. Richardson, M. Altaf Arain, Almut Arneth, Christian Bernhofer, Damien Bonal, Jiquan Chen, Damiano Gianelle, Nadine Gobron, Gerald Kiely, Werner Kutsch, Gitta Lasslop, Beverly E. Law, Anders Lindroth, Lutz Merbold, Leonardo Montagnani, Eddy J. Moors, Dario Papale, Matteo Sottocornola, Francesco Vaccari, and Christopher Williams.

- Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. *J. Geophys. Res.*, 116(G3):G00J07–, September 2011. ISSN 2156-2202. URL <http://dx.doi.org/10.1029/2010JG001566>.
- J. W. Kaiser, A. Heil, M. O. Andreae, A. Benedetti, N. Chubarova, L. Jones, J.-J. Morcrette, M. Razinger, M. G. Schultz, M. Suttie, and G. R. van der Werf. Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. *Biogeosciences*, 9(1):527–554, 2012. doi: 10.5194/bg-9-527-2012. URL <http://www.biogeosciences.net/9/527/2012/>.
- Thomas Kaminski, Peter J. Rayner, Martin Heimann, and Ian G. Enting. On aggregation errors in atmospheric transport inversions. *Journal of Geophysical Research: Atmospheres*, 106(D5):4703–4715, 2001. ISSN 2156-2202. doi: 10.1029/2000JD900581. URL <http://dx.doi.org/10.1029/2000JD900581>.
- U. Karstens, M. Gloor, M. Heimann, and C. Rödenbeck. Insights from simulations with high-resolution transport and process models on sampling of the atmosphere for constraining midlatitude land carbon sinks. *Journal of Geophysical Research: Atmospheres*, 111(D12):n/a–n/a, 2006. ISSN 2156-2202. doi: 10.1029/2005JD006278. URL <http://dx.doi.org/10.1029/2005JD006278>.
- Charles D. Keeling. The concentration and isotopic abundances of carbon dioxide in the atmosphere. *Tellus*, 12(2):200–203, 1960. ISSN 2153-3490. doi: 10.1111/j.2153-3490.1960.tb01300.x. URL <http://dx.doi.org/10.1111/j.2153-3490.1960.tb01300.x>.
- G. Krinner, Nicolas Viovy, Nathalie de Noblet-Ducoudré, Jérôme Ogée, Jan Polcher, Pierre Friedlingstein, Philippe Ciais, Stephen Sitch, and I. Colin Prentice. A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system. *Global Biogeochem. Cycles*, 19(1):GB1015–, March 2005. ISSN 1944-9224. URL <http://dx.doi.org/10.1029/2003GB002199>.
- M. Krol, S. Houweling, B. Bregman, M. van den Broek, A. Segers, P. van Velthoven, W. Peters, F. Dentener, and P. Bergamaschi. The two-way nested global chemistry-transport zoom model tm5: algorithm and applications. *Atmospheric Chemistry and Physics*, 5(2):417–432, 2005. doi: 10.5194/acp-5-417-2005. URL <http://www.atmos-chem-phys.net/5/417/2005/>.
- S. S. Kulawik, D. B. A. Jones, R. Nassar, F. W. Irion, J. R. Worden, K. W. Bowman, T. Machida, H. Matsueda, Y. Sawa, S. C. Biraud, M. L. Fischer, and A. R. Jacobson. Characterization of tropospheric emission spectrometer (tes) co<sub>2</sub> for carbon cycle science. *Atmospheric Chemistry and Physics*, 10(12):5601–5623, 2010. doi: 10.5194/acp-10-5601-2010. URL <http://www.atmos-chem-phys.net/10/5601/2010/>.
- Akihiko Kuze, Hiroshi Suto, Masakatsu Nakajima, and Takashi Hamazaki. Thermal and near infrared sensor for carbon observation fourier-transform spectrometer on the greenhouse gases observing satellite for greenhouse gases monitoring. *Appl. Opt.*, 48(35):6716–6733, Dec 2009. doi: 10.1364/AO.48.006716. URL <http://ao.osa.org/abstract.cfm?URI=ao-48-35-6716>.

- Peter Landschützer, Nicolas Gruber, F. Alexander Haumann, Christian Rödenbeck, Dorothee C. E. Bakker, Steven van Heuven, Mario Hoppema, Nicolas Metzl, Colm Sweeney, Taro Takahashi, Bronte Tilbrook, and Rik Wanninkhof. The reinvigoration of the southern ocean carbon sink. *Science*, 349(6253):1221–1224, September 2015. URL <http://science.sciencemag.org/content/349/6253/1221.abstract>.
- T. Lauvaux, M. Uliasz, C. Sarrat, F. Chevallier, P. Bousquet, C. Lac, K. J. Davis, P. Ciais, A. S. Denning, and P. J. Rayner. Mesoscale inversion: first results from the ceres campaign with synthetic data. *Atmospheric Chemistry and Physics*, 8(13):3459–3471, 2008. doi: 10.5194/acp-8-3459-2008. URL <http://www.atmos-chem-phys.net/8/3459/2008/>.
- R. M. Law, W. Peters, C. Rödenbeck, C. Aulagnier, I. Baker, D. J. Bergmann, P. Bousquet, J. Brandt, L. Bruhwiler, P. J. Cameron-Smith, J. H. Christensen, F. Delage, A. S. Denning, S. Fan, C. Geels, S. Houweling, R. Imasu, U. Karstens, S. R. Kawa, J. Kleist, M. C. Krol, S.-J. Lin, R. Lokupitiya, T. Maki, S. Maksyutov, Y. Niwa, R. Onishi, N. Parazoo, P. K. Patra, G. Pieterse, L. Rivier, M. Satoh, S. Serrar, S. Taguchi, M. Takigawa, R. Vautard, A. T. Vermeulen, and Z. Zhu. TransCom model simulations of hourly atmospheric CO<sub>2</sub>: Experimental overview and diurnal cycle results for 2002. *Global Biogeochemical Cycles*, 22(3):n/a–n/a, 2008. ISSN 1944-9224. doi: 10.1029/2007GB003050. URL <http://dx.doi.org/10.1029/2007GB003050>. GB3009.
- Jung-Eun Lee, Christian Frankenberg, Christiaan van der Tol, Joseph A. Berry, Luis Guanter, C. Kevin Boyce, Joshua B. Fisher, Eric Morrow, John R. Worden, Salvi Asefi, Grayson Badgley, and Sassan Saatchi. Forest productivity and water stress in amazonia: observations from gosat chlorophyll fluorescence. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1761), 2013. ISSN 0962-8452. doi: 10.1098/rspb.2013.0171. URL <http://rspb.royalsocietypublishing.org/content/280/1761/20130171>.
- J. Lloyd, O. Kolle, H. Fritsch, S. R. de Freitas, M. A. F. Silva Dias, P. Artaxo, A. D. Nobre, A. C. de Araújo, B. Kruijt, L. Sogacheva, G. Fisch, A. Thielmann, U. Kuhn, and M. O. Andreae. An airborne regional carbon balance for Central Amazonia. *Biogeosciences*, 4(5):759–768, 2007. doi: 10.5194/bg-4-759-2007. URL <http://www.biogeosciences.net/4/759/2007/>.
- K. M. Longo, S. R. Freitas, M. Pirre, V. Marécal, L. F. Rodrigues, J. Panetta, M. F. Alonso, N. E. Rosário, D. S. Moreira, M. S. Gácita, J. Arteta, R. Fonseca, R. Stockler, D. M. Katsurayama, A. Fazenda, and M. Bela. The chemistry catts-brams model (ccatts-brams 4.5): a regional atmospheric model system for integrated air quality and weather forecasting and research. *Geoscientific Model Development*, 6(5):1389–1405, 2013. doi: 10.5194/gmd-6-1389-2013. URL <http://www.geosci-model-dev.net/6/1389/2013/>.
- F. Maignan, F.-M. Bréon, F. Chevallier, N. Viovy, P. Ciais, C. Garrec, J. Trues, and M. Mancip. Evaluation of a global vegetation model using time series of satellite vegetation indices. *Geoscientific Model Development*, 4(4):1103–1114, 2011. doi: 10.5194/gmd-4-1103-2011. URL <http://www.geosci-model-dev.net/4/1103/2011/>.
- Yadvinder Malhi, Timothy R. Baker, Oliver L. Phillips, Samuel Almeida, Esteban Alvarez, Luzmilla Arroyo, Jerome Chave, Claudia I. Czimczik, Anthony Di Fiore, Niro Higuchi, Timothy J. Killeen, Susan G. Laurance, William F. Laurance, Simon L. Lewis, Lina

- María Mercado Montoya, Abel Monteagudo, David A. Neill, Percy Núñez Vargas, Sandra Patiño, Nigel C.A. Pitman, Carlos Alberto Quesada, Rafael Salomão, José Natalino Macedo Silva, Armando Torres Lezama, Rodolfo Vásquez Martínez, John Terborgh, Barbara Vinceti, and Jon Lloyd. The above-ground coarse wood productivity of 104 neotropical forest plots. *Global Change Biology*, 10(5):563–591, 2004. ISSN 1365-2486. doi: 10.1111/j.1529-8817.2003.00778.x. URL <http://dx.doi.org/10.1111/j.1529-8817.2003.00778.x>.
- Jose A. Marengo, Javier Tomasella, Lincoln M. Alves, Wagner R. Soares, and Daniel A. Rodriguez. The drought of 2010 in the context of historical droughts in the amazon region. *Geophysical Research Letters*, 38(12):n/a–n/a, 2011. ISSN 1944-8007. doi: 10.1029/2011GL047436. URL <http://dx.doi.org/10.1029/2011GL047436>. L12703.
- Belinda E. Medlyn, Sonke Zaehle, Martin G. De Kauwe, Anthony P. Walker, Michael C. Dietze, Paul J. Hanson, Thomas Hickler, Atul K. Jain, Yiqi Luo, William Parton, I. Colin Prentice, Peter E. Thornton, Shusen Wang, Ying-Ping Wang, Ensheng Weng, Colleen M. Iversen, Heather R. McCarthy, Jeffrey M. Warren, Ram Oren, and Richard J. Norby. Using ecosystem experiments to improve vegetation models. *Nature Climate Change*, 5(6):528–534, June 2015. ISSN 1758-678X. URL <http://dx.doi.org/10.1038/nclimate2621>.
- L. Menut, B. Bessagnet, D. Khvorostyanov, M. Beekmann, N. Blond, A. Colette, I. Coll, G. Curci, G. Foret, A. Hodzic, S. Mailler, F. Meleux, J.-L. Monge, I. Pison, G. Siour, S. Turquety, M. Valari, R. Vautard, and M. G. Vivanco. Chimere 2013: a model for regional atmospheric composition modelling. *Geoscientific Model Development*, 6(4):981–1028, 2013. doi: 10.5194/gmd-6-981-2013. URL <http://www.geosci-model-dev.net/6/981/2013/>.
- L. Molina, G. Broquet, P. Imbach, F. Chevallier, B. Poulter, D. Bonal, B. Burban, M. Ramonet, L. V. Gatti, S. C. Wofsy, J. W. Munger, E. Dlugokencky, and P. Ciais. On the ability of a global atmospheric inversion to constrain variations of CO<sub>2</sub> fluxes over amazonia. *Atmos. Chem. Phys.*, 15(14):8423–8438, 2015. doi: 10.5194/acp-15-8423-2015. URL <http://www.atmos-chem-phys.net/15/8423/2015/>.
- D. S. Moreira, S. R. Freitas, J. P. Bonatti, L. M. Mercado, N. M. É. Rosário, K. M. Longo, J. B. Miller, M. Gloor, and L. V. Gatti. Coupling between the JULES land-surface scheme and the CCATT-BRAMS atmospheric chemistry model (JULES-CCATT-BRAMS1.0): applications to numerical weather forecasting and the CO<sub>2</sub> budget in South America. *Geoscientific Model Development*, 6(4):1243–1259, 2013. doi: 10.5194/gmd-6-1243-2013. URL <http://www.geosci-model-dev.net/6/1243/2013/>.
- R. Nassar, D. B. A. Jones, S. S. Kulawik, J. R. Worden, K. W. Bowman, R. J. Andres, P. Suntharalingam, J. M. Chen, C. A. M. Brenninkmeijer, T. J. Schuck, T. J. Conway, and D. E. Worthy. Inverse modeling of CO<sub>2</sub> sources and sinks using satellite observations of CO<sub>2</sub> from TES and surface flask measurements. *Atmospheric Chemistry and Physics*, 11(12):6029–6047, 2011. doi: 10.5194/acp-11-6029-2011. URL <http://www.atmos-chem-phys.net/11/6029/2011/>.
- Jos G. J. Olivier, John A. Van Aardenne, Frank J. Dentener, Valerio Pagliari, Laurens N. Ganzeveld, and Jeroen A. H. W. Peters. Recent trends in global greenhouse gas emissions: regional trends 1970–2000 and spatial distribution of key sources in 2000. *En-*

- Environmental Sciences*, 2(2-3):81–99, 2005. doi: 10.1080/15693430500400345. URL <http://dx.doi.org/10.1080/15693430500400345>.
- N. C. Parazoo, A. S. Denning, S. R. Kawa, K. D. Corbin, R. S. Lokupitiya, and I. T. Baker. Mechanisms for synoptic variations of atmospheric  $\text{CO}_2$  in North America, South America and Europe. *Atmos. Chem. Phys.*, 8(23):7239–7254, 2008. doi: 10.5194/acp-8-7239-2008. URL <http://www.atmos-chem-phys.net/8/7239/2008/>.
- Nicholas C. Parazoo, Kevin Bowman, Christian Frankenberg, Jung-Eun Lee, Joshua B. Fisher, John Worden, Dylan B. A. Jones, Joseph Berry, G. James Collatz, Ian T. Baker, Martin Jung, Junjie Liu, Gregory Osterman, Chris O’Dell, Athena Sparks, Andre Butz, Sandrine Guerlet, Yukio Yoshida, Huilin Chen, and Christoph Gerbig. Interpreting seasonal changes in the carbon balance of southern Amazonia using measurements of  $\text{XCO}_2$  and chlorophyll fluorescence from GOSAT. *Geophys. Res. Lett.*, 40(11):2829–2833, June 2013. ISSN 1944-8007. URL <http://dx.doi.org/10.1002/grl.50452>.
- W. Peters, J. B. Miller, J. Whitaker, A. S. Denning, A. Hirsch, M. C. Krol, D. Zupanski, L. Bruhwiler, and P. P. Tans. An ensemble data assimilation system to estimate  $\text{CO}_2$  surface fluxes from atmospheric trace gas observations. *Journal of Geophysical Research: Atmospheres*, 110(D24):n/a–n/a, 2005. ISSN 2156-2202. doi: 10.1029/2005JD006157. URL <http://dx.doi.org/10.1029/2005JD006157>. D24304.
- Wouter Peters, Andrew R. Jacobson, Colm Sweeney, Arlyn E. Andrews, Thomas J. Conway, Kenneth Masarie, John B. Miller, Lori M. P. Bruhwiler, Gabrielle Pétron, Adam I. Hirsch, Douglas E. J. Worthy, Guido R. van der Werf, James T. Randerson, Paul O. Wennberg, Maarten C. Krol, , and Pieter P. Tans. An atmospheric perspective on north american carbon dioxide exchange: Carbontracker. *Proceedings of the National Academy of Sciences*, 104(48):18925–18930, November 2007. doi: 10.1073/pnas.0708986104.
- P. Peylin, R. M. Law, K. R. Gurney, F. Chevallier, A. R. Jacobson, T. Maki, Y. Niwa, P. K. Patra, W. Peters, P. J. Rayner, C. Rödenbeck, I. T. van der Laan-Luijkx, and X. Zhang. Global atmospheric carbon budget: results from an ensemble of atmospheric  $\text{CO}_2$  inversions. *Biogeosciences*, 10(10):6699–6720, 2013. doi: 10.5194/bg-10-6699-2013. URL <http://www.biogeosciences.net/10/6699/2013/>.
- Philippe Peylin, David Baker, Jorge Sarmiento, Philippe Ciais, and Philippe Bousquet. Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric  $\text{CO}_2$  data. *J. Geophys. Res.*, 107(D19):4385–, October 2002. ISSN 2156-2202. URL <http://dx.doi.org/10.1029/2001JD000857>.
- Oliver L. Phillips, Luiz E. O. C. Aragão, Simon L. Lewis, Joshua B. Fisher, Jon Lloyd, Gabriela López-González, Yadvinder Malhi, Abel Monteagudo, Julie Peacock, Carlos A. Quesada, Geertje van der Heijden, Samuel Almeida, Iêda Amaral, Luzmila Arroyo, Gerardo Aymard, Tim R. Baker, Olaf Bánki, Lilian Blanc, Damien Bonal, Paulo Brando, Jerome Chave, Atila Cristina Alves de Oliveira, Nallaret Dávila Cardozo, Claudia I. Czimczik, Ted R. Feldpausch, Maria Aparecida Freitas, Emanuel Gloor, Niro Higuchi, Eliana Jiménez, Gareth Lloyd, Patrick Meir, Casimiro Mendoza, Alexandra Morel, David A. Neill, Daniel Nepstad, Sandra Patiño, Maria Cristina Peñuela, Adriana Prieto, Fredy Ramírez, Michael Schwarz, Javier Silva, Marcos Silveira, Anne Sota Thomas, Hans ter Steege, Juliana Stropp, Rodolfo Vásquez, Przemyslaw Zelazowski, Esteban Alvarez Dávila, Sandy

- Andelman, Ana Andrade, Kuo-Jung Chao, Terry Erwin, Anthony Di Fiore, Eurídice Honorio C., Helen Keeling, Tim J. Killeen, William F. Laurance, Antonio Peña Cruz, Nigel C. A. Pitman, Percy Núñez Vargas, Hirma Ramírez-Angulo, Agustín Rudas, Rafael Salamão, Natalino Silva, John Terborgh, and Armando Torres-Lezama. Drought Sensitivity of the Amazon Rainforest. *Science*, 323(5919):1344–1347, March 2009. URL <http://www.sciencemag.org/content/323/5919/1344.abstract>.
- Vânia R. Pivello. The Use of Fire in the Cerrado and Amazonian Rainforests of Brazil: Past and Present. *The Journal of the Association for Fire Ecology*, 7(1):24–39, 2011. doi: 10.4996/fireecology.0701024. URL <http://dx.doi.org/10.4996/fireecology.0701024>.
- J. Pongratz, C. H. Reick, T. Raddatz, and M. Claussen. Effects of anthropogenic land cover change on the carbon cycle of the last millennium. *Global Biogeochemical Cycles*, 23(4):n/a–n/a, 2009. ISSN 1944-9224. doi: 10.1029/2009GB003488. URL <http://dx.doi.org/10.1029/2009GB003488>. GB4001.
- C. A. Quesada, O. L. Phillips, M. Schwarz, C. I. Czimczik, T. R. Baker, S. Patiño, N. M. Fyllas, M. G. Hodnett, R. Herrera, S. Almeida, E. Alvarez Dávila, A. Arneeth, L. Arroyo, K. J. Chao, N. Dezzeo, T. Erwin, A. di Fiore, N. Higuchi, E. Honorio Coronado, E. M. Jimenez, T. Killeen, A. T. Lezama, G. Lloyd, G. López-González, F. J. Luizão, Y. Malhi, A. Monteagudo, D. A. Neill, P. Núñez Vargas, R. Paiva, J. Peacock, M. C. Peñuela, A. Peña Cruz, N. Pitman, N. Priante Filho, A. Prieto, H. Ramírez, A. Rudas, R. Salomão, A. J. B. Santos, J. Schmerler, N. Silva, M. Silveira, R. Vásquez, I. Vieira, J. Terborgh, and J. Lloyd. Basin-wide variations in amazon forest structure and function are mediated by both soils and climate. *Biogeosciences*, 9(6):2203–2246, 2012. doi: 10.5194/bg-9-2203-2012. URL <http://www.biogeosciences.net/9/2203/2012/>.
- Barroso Mário Ramos-Neto and Regina Vânia Pivello. Lightning fires in a brazilian savanna national park: Rethinking management strategies. *Environmental Management*, 26(6): 675–684, 2000. ISSN 1432-1009. doi: 10.1007/s002670010124. URL <http://dx.doi.org/10.1007/s002670010124>.
- J T Randerson, G R van der Werf, L Giglio, G J Collatz, and P S Kasibhatla. Global {F}ire {E}missions {D}atabase, {V}ersion 3 ({GFED}v3.1), 2013. URL <http://dx.doi.org/10.3334/ORNLDAAC/1191>.
- Natalia Restrepo-Coupe, Humberto R. da Rocha, Lucy R. Hutyrá, Alessandro C. da Araujo, Laura S. Borma, Bradley Christoffersen, Osvaldo M.R. Cabral, Plinio B. de Camargo, Fernando L. Cardoso, Antonio C. Lola da Costa, David R. Fitzjarrald, Michael L. Goulden, Bart Kruijt, Jair M.F. Maia, Yadvinder S. Malhi, Antonio O. Manzi, Scott D. Miller, Antonio D. Nobre, Celso von Randow, Leonardo D. Abreu Sá, Ricardo K. Sakai, Julio Tota, Steven C. Wofsy, Fabricio B. Zanchi, and Scott R. Saleska. What drives the seasonality of photosynthesis across the amazon basin? a cross-site analysis of eddy flux tower measurements from the brasil flux network. *Agricultural and Forest Meteorology*, 182–183(0):128 – 144, 2013. ISSN 0168-1923. doi: <http://dx.doi.org/10.1016/j.agrformet.2013.04.031>. URL <http://www.sciencedirect.com/science/article/pii/S0168192313001184>.
- M. Reuter, M. Buchwitz, M. Hilker, J. Heymann, O. Schneising, D. Pillai, H. Bovensmann, J. P. Burrows, H. Bösch, R. Parker, A. Butz, O. Hasekamp, C. W. O’Dell, Y. Yoshida, C. Gerbig, T. Nehr Korn, N. M. Deutscher, T. Warneke, J. Notholt, F. Hase, R. Kivi, R. Sussmann,

- T. Machida, H. Matsueda, and Y. Sawa. Satellite-inferred european carbon sink larger than expected. *Atmospheric Chemistry and Physics*, 14(24):13739–13753, 2014. doi: 10.5194/acp-14-13739-2014. URL <http://www.atmos-chem-phys.net/14/13739/2014/>.
- Richard W. Reynolds, Nick A. Rayner, Thomas M. Smith, Diane C. Stokes, and Wanqiu Wang. An improved in situ and satellite sst analysis for climate. *J. Climate*, 15(13):1609–1625, 2002. ISSN 0894-8755. doi: 10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2. URL [http://dx.doi.org/10.1175/1520-0442\(2002\)015<1609:AIISAS>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2002)015<1609:AIISAS>2.0.CO;2).
- C. Rödenbeck, S. Houweling, M. Gloor, and M. Heimann. CO<sub>2</sub> flux history 1982-2001 inferred from atmospheric data using a global inversion of atmospheric transport. *Atmospheric Chemistry and Physics*, 3(6):1919–1964, 2003. doi: 10.5194/acp-3-1919-2003. URL <http://www.atmos-chem-phys.net/3/1919/2003/>.
- C. Rödenbeck, C. Gerbig, K. Trusilova, and M. Heimann. A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models. *Atmospheric Chemistry and Physics*, 9(14):5331–5342, 2009. doi: 10.5194/acp-9-5331-2009. URL <http://www.atmos-chem-phys.net/9/5331/2009/>.
- C. Rödenbeck, D. C. E. Bakker, N. Gruber, Y. Iida, A. R. Jacobson, S. Jones, P. Landschützer, N. Metzl, S. Nakaoka, A. Olsen, G.-H. Park, P. Peylin, K. B. Rodgers, T. P. Sasse, U. Schuster, J. D. Shutler, V. Valsala, R. Wanninkhof, and J. Zeng. Data-based estimates of the ocean carbon sink variability – first results of the surface ocean *p*<sub>co2</sub> mapping intercomparison (socom). *Biogeosciences*, 12(23):7251–7278, 2015. doi: 10.5194/bg-12-7251-2015. URL <http://www.biogeosciences.net/12/7251/2015/>.
- Clive D. Rodgers. *Inverse Methods For Atmospheric Sounding: Theory And Practice*. World Scientific, 2000.
- N. E. Rosário, K. M. Longo, S. R. Freitas, M. A. Yamasoe, and R. M. Fonseca. Modeling the South American regional smoke plume: aerosol optical depth variability and surface shortwave flux perturbation. *Atmospheric Chemistry and Physics*, 13(6):2923–2938, 2013. doi: 10.5194/acp-13-2923-2013. URL <http://www.atmos-chem-phys.net/13/2923/2013/>.
- L. Rossato, R. C. S. Alvalá, and J. Tomasella. Distribuição geográfica da capacidade de armazenamento de água e das propriedades física do solo no Brasil. *X Congresso Brasileiro de Meteorologia / VIII Congresso da FLISMET*, 1998.
- Susanna Rutledge, David I. Campbell, Dennis Baldocchi, and Louis A. Schipper. Photodegradation leads to increased carbon dioxide losses from terrestrial organic matter. *Global Change Biology*, 16(11):3065–3074, 2010. ISSN 1365-2486. doi: 10.1111/j.1365-2486.2009.02149.x. URL <http://dx.doi.org/10.1111/j.1365-2486.2009.02149.x>.
- C. Rödenbeck, S. Houweling, M. Gloor, and M. Heimann. Time-dependent atmospheric CO<sub>2</sub> inversions based on interannually varying tracer transport. *Tellus B*, 55(2), 2003. ISSN 1600-0889. URL <http://www.tellusb.net/index.php/tellusb/article/view/16707>.

- Sassan Saatchi, Salvi Asefi-Najafabady, Yadvinder Malhi, Luiz E. O. C. Aragão, Liana O. Anderson, Ranga B. Myneni, and Ramakrishna Nemani. Persistent effects of a severe drought on Amazonian forest canopy. *Proc. Natl. Acad. Sci. U.S.A.*, 110:565–570, December 2012. URL <http://www.pnas.org/content/early/2012/12/19/1204651110.abstract>.
- Scott R. Saleska, Scott D. Miller, Daniel M. Matross, Michael L. Goulden, Steven C. Wofsy, Humberto R. da Rocha, Plinio B. de Camargo, Patrick Crill, Bruce C. Daube, Helber C. de Freitas, Lucy Hutyla, Michael Keller, Volker Kirchhoff, Mary Menton, J. William Munger, Elizabeth Hammond Pyle, Amy H. Rice, and Hudson Silva. Carbon in Amazon Forests: Unexpected Seasonal Fluxes and Disturbance-Induced Losses. *Science*, 302(5650):1554–1557, November 2003. URL <http://www.sciencemag.org/content/302/5650/1554.abstract>.
- Scott R. Saleska, Kamel Didan, Alfredo R. Huete, and Humberto R. da Rocha. Amazon Forests Green-Up During 2005 Drought. *Science*, 318(5850):612–612, October 2007. URL <http://www.sciencemag.org/content/318/5850/612.abstract>.
- Arindam Samanta, Sangram Ganguly, Hirofumi Hashimoto, Sadashiva Devadiga, Eric Vermote, Yuri Knyazikhin, Ramakrishna R. Nemani, and Ranga B. Myneni. Amazon forests did not green-up during the 2005 drought. *Geophys. Res. Lett.*, 37(5):L05401–, March 2010. ISSN 1944-8007. URL <http://dx.doi.org/10.1029/2009GL042154>.
- M. F. Sestini, E. S. Reimer, D. M. Valeriano, R. C. S. Alvalá, E. M. K. Mello, C. S. Chan, , and C. A. Nobre. Mapa de cobertura da terra da Amazônia legal para uso em modelos meteorológicos. *Simpósio Brasileiro de Sensoriamento Remoto*, 11:2901–2906, 2003.
- S. Sitch, P. Friedlingstein, N. Gruber, S. D. Jones, G. Murray-Tortarolo, A. Ahlström, S. C. Doney, H. Graven, C. Heinze, C. Huntingford, S. Levis, P. E. Levy, M. Lomas, B. Poulter, N. Viovy, S. Zaehle, N. Zeng, A. Arneth, G. Bonan, L. Bopp, J. G. Canadell, F. Chevallier, P. Ciais, R. Ellis, M. Gloor, P. Peylin, S. L. Piao, C. Le Quéré, B. Smith, Z. Zhu, and R. Myneni. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12(3):653–679, 2015. doi: 10.5194/bg-12-653-2015. URL <http://www.biogeosciences.net/12/653/2015/>.
- Xiao-Peng Song, Chengquan Huang, Sassan S. Saatchi, Matthew C. Hansen, and John R. Townshend. Annual carbon emissions from deforestation in the amazon basin between 2000 and 2010. *PLoS ONE*, 10(5):1–21, 05 2015. doi: 10.1371/journal.pone.0126754. URL <http://dx.doi.org/10.1371%2Fjournal.pone.0126754>.
- J. Staufer, G. Broquet, F.-M. Bréon, V. Puygrenier, F. Chevallier, I. Xueref-Rémy, E. Dieudonné, M. Lopez, M. Schmidt, M. Ramonet, O. Perrussel, C. Lac, L. Wu, and P. Ciais. A first year-long estimate of the paris region fossil fuel co2 emissions based on atmospheric inversion. *Atmospheric Chemistry and Physics Discussions*, 2016:1–34, 2016. doi: 10.5194/acp-2016-191. URL <http://www.atmos-chem-phys-discuss.net/acp-2016-191/>.
- Britton B. Stephens, Kevin R. Gurney, Pieter P. Tans, Colm Sweeney, Wouter Peters, Lori Bruhwiler, Philippe Ciais, Michel Ramonet, Philippe Bousquet, Takakiyo Nakazawa, Shuji Aoki, Toshinobu Machida, Gen Inoue, Nikolay Vinnichenko, Jon Lloyd, Armin Jordan, Martin Heimann, Olga Shibistova, Ray L. Langenfelds, L. Paul Steele, Roger J. Francey,

- and A. Scott Denning. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric  $\text{CO}_2$ . *Science*, 316(5832):1732–1735, 2007. doi: 10.1126/science.1137004. URL <http://www.sciencemag.org/content/316/5832/1732.abstract>.
- Hiroshi Takagi, Sander Houweling, Robert J. Andres, Dmitry Belikov, Andrey Bril, Hartmut Boesch, Andre Butz, Sandrine Guerlet, Otto Hasekamp, Shamil Maksyutov, Isamu Morino, Tomohiro Oda, Christopher W. O'Dell, Sergey Oshchepkov, Robert Parker, Makoto Saito, Osamu Uchino, Tatsuya Yokota, Yukio Yoshida, and Vinu Valsala. Influence of differences in current GOSAT XCO<sub>2</sub> retrievals on surface flux estimation. *Geophysical Research Letters*, 41(7):2598–2605, 2014. ISSN 1944-8007. doi: 10.1002/2013GL059174. URL <http://dx.doi.org/10.1002/2013GL059174>. 2013GL059174.
- T. Takahashi, Sutherlandm S.C., and A. Kozyr. Global ocean surface water partial pressure of  $\text{CO}_2$  database: Measurements performed during 1957-2015 (version 2015), 2016. URL [http://cdiac.ornl.gov/oceans/LDEO\\_Underway\\_Database/](http://cdiac.ornl.gov/oceans/LDEO_Underway_Database/).
- Taro Takahashi, Stewart C. Sutherland, Rik Wanninkhof, Colm Sweeney, Richard A. Feely, David W. Chipman, Burke Hales, Gernot Friederich, Francisco Chavez, Christopher Sabine, Andrew Watson, Dorothee C.E. Bakker, Ute Schuster, Nicolas Metzl, Hisayuki Yoshikawa-Inoue, Masao Ishii, Takashi Midorikawa, Yukihiro Nojiri, Arne Körtzinger, Tobias Steinhoff, Mario Hoppema, Jon Olafsson, Thorarinn S. Arnarson, Bronte Tilbrook, Truls Johannessen, Are Olsen, Richard Bellerby, C.S. Wong, Bruno Delille, N.R. Bates, and Hein J.W. de Baar. Climatological mean and decadal change in surface ocean  $\text{pCO}_2$ , and net sea-air  $\text{CO}_2$  flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(8-10):554–577, April 2009. ISSN 0967-0645. URL <http://www.sciencedirect.com/science/article/pii/S0967064508004311>.
- Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics, Philadelphia, 2005.
- Hanna Tuomisto, Kalle Ruokolainen, Risto Kalliola, Ari Linna, Walter Danjoy, and Zoila Rodriguez. Dissecting amazonian biodiversity. *Science*, 269(5220):63–66, 1995. ISSN 0036-8075. doi: 10.1126/science.269.5220.63. URL <http://science.sciencemag.org/content/269/5220/63>.
- I. T. van der Laan-Luijkx, I. R. van der Velde, M. C. Krol, L. V. Gatti, L. G. Domingues, C. S. C. Correia, J. B. Miller, M. Gloor, T. T. van Leeuwen, J. W. Kaiser, C. Wiedinmyer, S. Basu, C. Clerbaux, and W. Peters. Response of the amazon carbon balance to the 2010 drought derived with carbontracker south america. *Global Biogeochemical Cycles*, 29(7):1092–1108, 2015. ISSN 1944-9224. doi: 10.1002/2014GB005082. URL <http://dx.doi.org/10.1002/2014GB005082>. 2014GB005082.
- G. R. van der Werf, D. C. Morton, R. S. DeFries, J. G. J. Olivier, P. S. Kasibhatla, R. B. Jackson, G. J. Collatz, and J. T. Randerson.  $\text{CO}_2$  emissions from forest loss. *Nature Geoscience*, 2(11):737–738, November 2009. ISSN 1752-0894. URL <http://dx.doi.org/10.1038/ngeo671>.
- R. Wang, S. Tao, P. Ciais, H. Z. Shen, Y. Huang, H. Chen, G. F. Shen, B. Wang, W. Li, Y. Y. Zhang, Y. Lu, D. Zhu, Y. C. Chen, X. P. Liu, W. T. Wang, X. L. Wang, W. X. Liu, B. G. Li, and S. L. Piao. High-resolution mapping of combustion processes and implications for

- CO<sub>2</sub> emissions. *Atmospheric Chemistry and Physics*, 13(10):5189–5203, 2013. doi: 10.5194/acp-13-5189-2013. URL <http://www.atmos-chem-phys.net/13/5189/2013/>.
- C. Wiedinmyer, S. K. Akagi, R. J. Yokelson, L. K. Emmons, J. A. Al-Saadi, J. J. Orlando, and A. J. Soja. The fire inventory from near (finn): a high resolution global model to estimate the emissions from open burning. *Geoscientific Model Development*, 4(3):625–641, 2011. doi: 10.5194/gmd-4-625-2011. URL <http://www.geosci-model-dev.net/4/625/2011/>.
- Steven C. Wofsy, Robert C. Harriss, and Warren A. Kaplan. Carbon dioxide in the atmosphere over the amazon basin. *Journal of Geophysical Research: Atmospheres*, 93(D2):1377–1387, 1988. ISSN 2156-2202. doi: 10.1029/JD093iD02p01377. URL <http://dx.doi.org/10.1029/JD093iD02p01377>.
- L. Wu, G. Broquet, P. Ciais, V. Bellassen, F. Vogel, F. Chevallier, I. Xueref-Remy, and Y. Wang. What would dense atmospheric observation networks bring to the quantification of city co2 emissions? *Atmospheric Chemistry and Physics*, 16(12):7743–7771, 2016. doi: 10.5194/acp-16-7743-2016. URL <http://www.atmos-chem-phys.net/16/7743/2016/>.
- Liang Xu, Sassan S Saatchi, Yan Yang, Ranga B Myneni, Christian Frankenberg, Diya Chowdhury, and Jian Bi. Satellite observation of tropical forest seasonality: spatial patterns of carbon exchange in Amazonia. *Environmental Research Letters*, 10(8):084005, 2015. URL <http://stacks.iop.org/1748-9326/10/i=8/a=084005>.
- Z.-L. Yang, R. E. Dickinson, A. Henderson-Sellers, and A. J. Pitman. Preliminary study of spin-up processes in land surface models with the first stage data of Project for Intercomparison of Land Surface Parameterization Schemes Phase 1(a). *Journal of Geophysical Research: Atmospheres*, 100(D8):16553–16578, 1995. ISSN 2156-2202. doi: 10.1029/95JD01076. URL <http://dx.doi.org/10.1029/95JD01076>.
- Yi Yin, Philippe Ciais, Frederic Chevallier, Guido R. van der Werf, Thierry Fanin, Gregoire Broquet, Hartmut Boesch, Anne Cozic, Didier Hauglustaine, Sophie Szopa, and Yilong Wang. Variability of fire carbon emissions in equatorial asia and its nonlinear sensitivity to el niño. *Geophysical Research Letters*, 43(19):10,472–10,479, 2016. ISSN 1944-8007. doi: 10.1002/2016GL070971. URL <http://dx.doi.org/10.1002/2016GL070971>. 2016GL070971.
- L. Zobler. Global soil types, 1-degree grid (zobler). data set. available on-line [<http://www.daac.ornl.gov>] from oak ridge national laboratory distributed active archive center, 1999.