Regional CO_2 inversions with LUMIA, the Lund University Modular Inversion Algorithm, v1.0

Guillaume Monteil[1] and Marko Scholze[1]

[1]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

Correspondence: Guillaume Monteil (guillaume.monteil@nateko.lu.se)

Abstract. Atmospheric inversions are used to derive constraints on the net sources and sinks of CO_2 and other stable atmospheric tracers from their observed concentrations. The resolution and accuracy the fluxes can be estimated with depends, among other factors, on the quality and density of the observational coverage, on the precision and accuracy of the transport model used by the inversion to relate fluxes to observations, and on the adaptation of the statistical approach to the problem studied (Michalak et al., 2016).

In recent years, there has been an increasing demand from stakeholders for inversions at higher spatial resolution (country scale), in particular in the framework of the Paris agreement. This step up in resolution is in theory enabled by the growing availability of observations from surface in-situ networks (such as ICOS in Europe) and from remote sensing products (OCO-2, GOSAT-2). The increase in the resolution of inversions is also a necessary step to provide efficient feedback to the bottom-up modelling community (vegetation models, fossil fuel emission inventories, etc.). It however calls for new developments in the inverse models: diversification of the inversion approaches, shift from global to regional inversions, improvement in the computational efficiency, etc.

We developed the Lund University Modular Inversion Algorithm (LUMIA) as a tool to address some of these new topics. LUMIA is meant to become a platform for inverse modelling developments at Lund University. It aims at being a flexible, yet simple and easy to maintain set of tools that the modellers can combine to build inverse modelling experiments. It is in particular designed to be transport model agnostic, which should facilitate isolating the transport model errors from those introduced by the inversion setup itself.

This paper describes briefly the LUMIA framework: the motivations for building it, the development principles, current status and future prospects. Then a first LUMIA inversion setup is presented, to perform regional CO_2 inversions in Europe, using in-situ data from surface and tall tower observation sites. Since LUMIA doesn't come with its own transport model, the transport of fluxes is computed using a coupling between the Lagrangian FLEXPART transport model (high resolution foreground transport) and the global coarse resolution TM5 transport model (following the approach of Rödenbeck et al. (2009)). This particular coupling is new and therefore also described in this paper.

1 Introduction

The accumulation of greenhouse gases in the atmosphere is the main driver of climate change. The largest contribution of anthropogenic activities to global warming is through the release of fossil carbon (mainly as CO_2) to the atmosphere, but other human activities such as land use change (for agriculture, deforestation, etc.) also play a significant role. The climate forcings from this increased CO_2 concentration is likely to induce feedbacks through reactions of the terrestrial ecosystems and of the oceans (Stocker et al., 2013). Our capacity to correctly predict climate change, anticipate and mitigate its effects depends therefore largely on our capacity to model and predict the evolution of carbon exchanges between the atmosphere and other reservoirs.

For future climate simulations, the only available option is through "direct" (bottom-up) modelling of the different components of the biogeochemical cycles, i.e. using models (numerical or statistical) that simulate, as accurately as possible (given the precision requirements of the simulation), greenhouse gas fluxes to and from the atmosphere. For past periods, however, the "inverse" (top-down) approach is also possible, in which the greenhouse gas fluxes are diagnosed from their observed impact on atmospheric greenhouse gas concentrations.

Direct and inverse approaches are complementary, the former can provide detailed estimates of the spatial and temporal variability of the fluxes, but often with large uncertainties on the total fluxes Sitch et al. (2015) . On the contrary, inverse approaches provide robust estimates of total fluxes at large scales consistent with the observations (e.g. Gurney et al. (2002)), but with poor sensitivity to smaller scales (e.g. Peylin et al. (2013)).

An atmospheric inverse model typically couples an atmospheric transport model (which computes the relationships between fluxes and concentrations) with an inversion algorithm, whose task is to determine the most likely set of fluxes, within some prior constraints and given the information from an observation ensemble (in a Bayesian approach). In practice, inversions are complex codes, computationally heavy. The complexity arises in a large part from the necessity to combine large quantities of informations from sometimes very heterogeneous datasets (various types of observations, flux estimates, meteorological forcings, etc.). The computational weight depends largely on that of the underlying transport model, which usually needs to be ran a large number of times (iteratively or as an ensemble).

In recent years, the availability of observations has grown by orders of magnitude, with the deployment of high-density surface observation networks (such as the Integrated Carbon Observation System, ICOS, in Europe) and the fast developments in satellite retrievals of tropospheric greenhouse gas concentrations (GOSAT, OCO-2, etc.). Meanwhile, the demand for inversions is increasing, in particular from stakeholders such as regional, national or trans-national governments who are interested in country-scale inversions as a means of quantifying their carbon emissions, in connection with emission reduction targets as defined in the Paris agreement (Ciais et al., 2015).

This context puts strains on the existing inverse models. The larger availability of high quality data means that fluxes can be constrained at finer scales, but it also means that models of higher definition and precision must be used. The development of regional inversions (of varying scales) allows in theory an efficient usage of high resolution data while preserving a reasonable computational cost, but comes with specific challenges such as the need of more boundary conditions and the lack of options

for cross-validation when the resolution increases and the domain size shrinks. The demands from various stakeholders (policy makers, bottom-up modellers, medias, etc.) also call for developments in the inversion techniques, with for instance a more pronounced focus on the quantification of anthropogenic sources (Ciais et al., 2015) or the optimization of ecosystem models parameters instead of CO_2 fluxes in carbon cycle data assimilation systems (CCDAS) (Kaminski et al., 2013).

To enable such progress in the method and quality of the inversions, it is important to have a robust and flexible tool. The purpose of LUMIA (Lund University Modular Inversion Algorithm) is to be a development platform for top-down experiments. LUMIA was developed from the start as a model-agnostic inversion tool, with a clear isolation of the data stream between the transport model and the optimization algorithm in an interface module. One of the main aims is to eventually allow a better characterization of the uncertainty associated to the transport model. Strong emphasis was put on the usability (low barrier entry code for newcomers, high degree of modularity to allow users to build their experiments in a very flexible way) and sustainability of the code (small, easily replaceable one-tasked modules instead of large multi-option ones).

This paper presents the LUMIA inversion framework, as well as a first application of regional (European) CO_2 inversions for Europe. The inversions use in-situ CO_2 observations from European tall towers (now part of the ICOS network, see https://www.icos-ri.eu) and rely on a regional transport model based on a new coupling between the FLEXPART Lagrangian particle dispersion model (Seibert and Frank, 2004; Pisso et al., 2019) (foreground, high resolution transport) and TM5-4DVAR (Meirink et al., 2008; Basu et al., 2013) (background concentrations). The paper is organized as follows: First, Section 2 presents the LUMIA framework (general principles and architecture). Then Section 3 presents the specific inverse modelling setup used here (including the FLEXPART-TM5 coupling). Sections 4 and 5 present the results from two set of inversions (against synthetic and real observations). Finally, a short discussion summarises the main outcomes of the paper in Section 6.

2 The LUMIA framework

2.1 Theoretical background

The general principle of an atmospheric inversion is to determine the most likely estimate of a set of variables controlling the atmospheric content and distribution of a tracer (typically sources and sinks, but also initial or boundary conditions), given a set of observations of that tracer's distribution in the atmosphere. The link between the set of parameters to optimize (control vector $\mathbf{x}$, of dimension $n_{\mathbf{x}}$) and the observed concentrations (observation vector $\mathbf{y}$, of dimension $n_{\mathbf{y}}$) is established by a numerical model of the atmospheric transport (and of any other physical process relating the state and observation vectors):

$$\mathbf{y} + \varepsilon_y = H(\mathbf{x} + \varepsilon_x) + \varepsilon_H \tag{1}$$

The observation operator H includes the transport model itself, but also any additional steps needed to express $\mathbf{y}$ as a function of $\mathbf{x}$ (aggregation/disaggregation of flux components, accounting of boundary conditions and of non-optimized fluxes, etc.).

The error terms ε_y, ε_x and ε_H are respectively the observation error, the control vector error and the model error (see Section 3.3.1). In the simplest cases, the system can be solved for $\mathbf{x}$ analytically, but most often inversions use the Bayesian inference approach: the optimal control vector $\hat{\mathbf{x}}$ is the one that allows the best statistical compromise between fitting the observations and limiting the departure from a prior estimation of the control vector $\mathbf{x_b}$ (accounting for the (prescribed) uncertainties in the observations and the prior). Mathematically, this means finding the vector $\hat{\mathbf{x}}$ that minimizes a cost function $J(\mathbf{x})$ defined (in our case) as

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2}\left(\mathbf{x}-\mathbf{x_b}\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x_b}\right) \\ &\quad + \frac{1}{2}\sum_j \left(\mathbf{Hx}-\mathbf{y_j}\right)^T \mathbf{R}^{-1}\left(\mathbf{Hx}-\mathbf{y_j}\right) \\ &= J_{\mathbf{b}} + J_{obs} \end{aligned} \tag{2}$$

where the prior ($\mathbf{B}$) and observation ($\mathbf{R}$) error covariance matrices weigh the relative contributions to the cost function of each departure from each prior control variable x_b^i and from each observation y^j. The optimal control vector $\hat{\mathbf{x}}$ is solved for analytically (for small scale problems) or approximated step-wise (variational and ensemble approaches are most common (Rayner et al., 2018)).

An inversion system is therefore the combination of an observation operator (i.e. transport model, sampling operator, etc.), an inversion technique and a set of assumptions on the prior values of the variables to estimates, their uncertainties and the uncertainties of the observations. Each of these components introduces its own share of uncertainty, which makes the results harder to interpret: which feature of the solution is real, and which is introduced by e.g. the transport model, or incorrect assumptions on some uncertainties?

2.2 The lumia python package

The LUMIA system is designed with the aim to provide the modularity needed to quantify the impact of the inversion design choices on the inversion results themselves. The strict isolation of the transport model also enables the transport model and the inversion algorithm to evolve independently. On the other hand, the modularity does lead to an increase in the overall complexity of the code (due to the need to develop and maintain generic interfaces), which can end up being counterproductive if it limits the performances and/or usability of the system. We nonetheless believe that the benefit of a higher modularity outweighs the risks. The potential adverse effects can be mitigated by careful design choices. The code is distributed as a single python package, with the following structure (see also Figure 1):

- The *lumia* folder contains the `lumia` python library, which implements the basic components of the inversion such as data storage (control vector, fluxes, observations, uncertainty matrices) and functions (forward and adjoint transport, conversion functions between fluxes and control vector; cost function evaluation, etc.).

- The *transport* folder contains the code that was used to implement the TM5-FLEXPART transport model coupling, described further in Section 3.2.

- The *src* folder contains the FORTRAN source code for the conjugate gradient minimizer used in the example inversions (see Section 3). Replacing this external code by a native python equivalent is planned.

- The *doc* folder contains a documentation, mainly in the form of jupyter-notebooks, and example data and configuration files.

- The *GMDD* folder contains the scripts and configuration files used for producing the results presented further down in this manuscript.

The package can be installed using the standard 'pip' command, which installs `lumia` and `transport` as python modules, which can then be imported from any python script. The `lumia` module itself has a relatively flat hierarchy, which limits the risk that replacing or changing one component prevents the others from working. The implementation of alternative features is preferably carried out via the development of alternative classes, which allows each individual class to remain compact and easy to understand and maintain.

The `lumia` and `transport` modules and their submodules can be used totally independently from the inversion scripts that are provided in the scripts folder. This allows their use in different contexts, such as development, pre/post-processing of the inversion data or during the analysis of the results (and eventually this helps keeping the inversion scripts compact, as they need only to focus on the inversion itself). The scope of the `lumia` library is intentionally vague: it should permit easily constructing inversion experiments and is primarily designed for it, but our current design choices should not over-constrain the alternative use cases (such as e.g. forward transport model experiments or optimization of land surface model parameters).

3 Test inversion setup

Our test inversion setup is designed to optimize the monthly net atmosphere-ecosystem carbon flux (NEE, Net Ecosystem Exchange) over Europe at a target horizontal resolution of $0.5°$, using CO_2 observations from the European ICOS network (or similar/precursor sites). Two series of inversions are presented: First, a series of Observing System Synthetic Experiments (OSSEs), using known truth and synthetic observations; then a series of inversions constrained by real observations. All the inversions are performed on a domain ranging from 15°W, 33°N to 35°E, 73°N (illustrated in Figure 2, and hereafter referred to as the Regional Inversion Domain, RID) and cover the year 2011. The following sections describe the variational inversion technique, the transport model and the problem constraints (prior fluxes and observations).

3.1 Inversion approach

We use a Bayesian variational inversion algorithm, similar to that used in TM5-4DVAR inversions (Basu et al., 2013; Meirink et al., 2008). In a variational inversion, the minimum of the cost function $J(\mathbf{x})$ (Equation 2) is solved for iteratively:

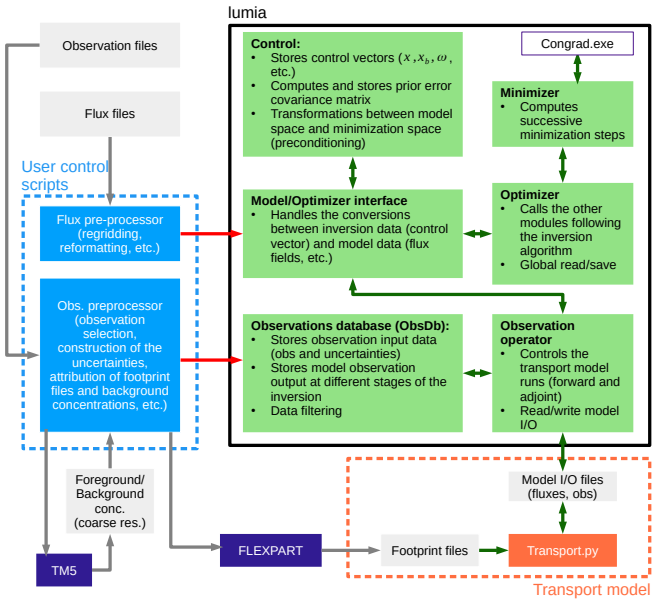

Figure 1. Inversion flow diagram. The green boxes represent code that is part of the lumia python module; the orange box shows operations performed by the atmospheric transport model (in our case a simple python script that reads in observations, fluxes and footprints, but a full transport model could be plugged in instead); the blue boxes show code that is typically user- and application-specific (pre-processing of data and main inversion control script). The boxes in grey mark on-disk data, and the boxes in purple show external executables

- An initial "prior" run is performed to compute the concentrations ($\mathbf{y_m} = \mathbf{H x_b}$) corresponding to the prior control vector $\mathbf{x_b}$ (since the transport is linear, $H(\mathbf{x}) \equiv \mathbf{H x}$).
- The local cost function ($J(\mathbf{x} = \mathbf{x_b})$) and cost function gradient ($\nabla_\mathbf{x} J(\mathbf{x} = \mathbf{x_b})$) are computed.
- A control vector increment ($\delta \mathbf{x}$) is deduced from the gradient, and the process is repeated from step 1 (with $\mathbf{x} = \mathbf{x_b} + \delta \mathbf{x}$), until a convergence criterion is reached.

The control vector increments are computed using an external library implementing the Lanczos algorithm (Lanczos, 1950). For efficiency (reduction of the number of iterations) and practicality (reduction of the number of large matrix multiplications) reasons, the optimization is performed on the preconditioned variable $\omega = \mathbf{B^{-1/2}}(\mathbf{x} - \mathbf{x_b})$ (following Courtier et al. (1994) and similar to the implementation in Basu et al. (2013). Equation 2 then becomes

$$J(\omega) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\left(\mathbf{H}\mathbf{B}^{1/2}\omega + \mathbf{d_0}\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\mathbf{B}^{1/2}\omega + \mathbf{d_0}\right) \tag{3}$$

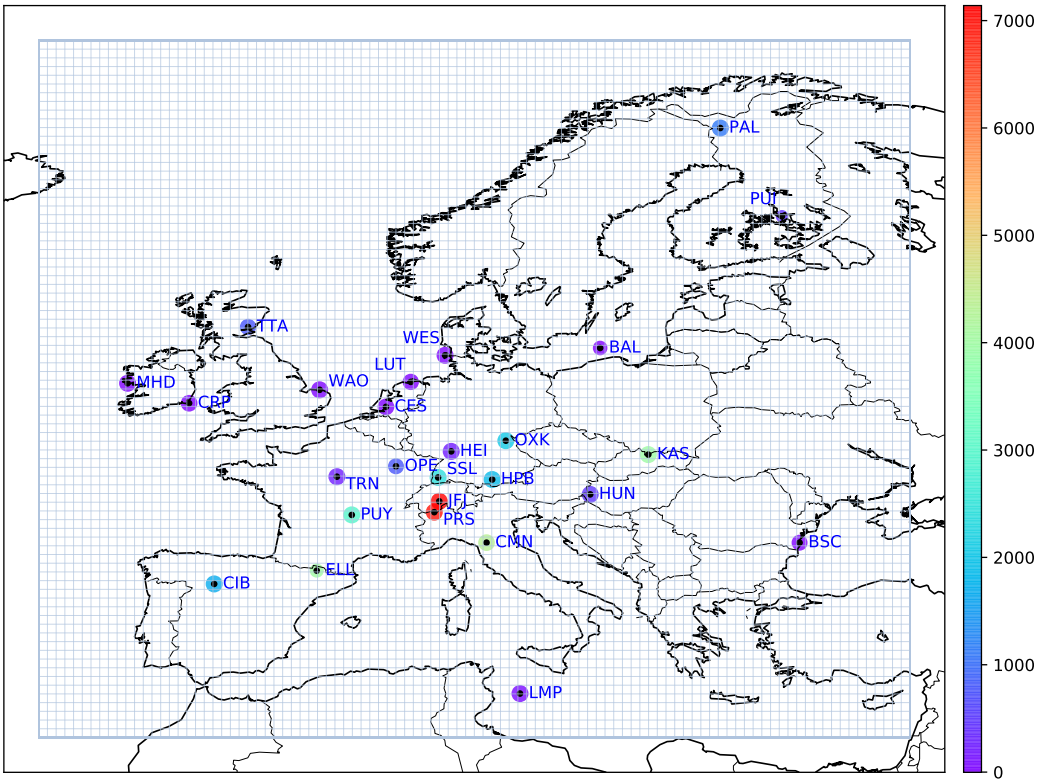

Figure 2. Regional inversion domain and location of the observation sites. The area of the dots is proportional to the number of observations available at each site (the actual number of observations is reduced by the filtering described in Section 3.3.4) and their color represents the altitude of the sites.

with $\mathbf{d_0} = \mathbf{Hx_b} - \mathbf{y}$ the prior model-data mismatches. In this formulation, the cost function gradient is given by

$$\begin{aligned}\nabla_\omega J &= \omega + \mathbf{B^{T/2}H^T R^{-1}}\left(\mathbf{Hx} - \mathbf{y}\right) \\ &= \omega + \mathbf{B^{T/2}}\nabla_\mathbf{x} J_{obs}\end{aligned} \tag{4}$$

The non-preconditioned observational cost function gradient $\nabla_x J_{obs} = \mathbf{H^T R^{-1}}\left(\mathbf{Hx} - \mathbf{y}\right)$ is computed using the adjoint technique (Errico, 1997). The transformation matrix $\mathbf{B^{1/2}}$ is obtained by eigen-value decomposition of $\mathbf{B}$. Note that in this formulation, the inverse of $\mathbf{B}$ (or the square root of its inverse) is actually never needed, making it possible to constrain the inversion with a non invertible matrix. In practice, the preconditioning adds two extra steps to the algorithm described above: conversion from ω to $\mathbf{x}$ ($\mathbf{x} = \mathbf{B^{1/2}}\omega + \mathbf{x_b}$) before applying the transport operator (i.e. running the transport model), just before

step 1; Conversion from $\nabla_{\mathbf{x}} J_{obs}$ to $\nabla_{\omega} J_{obs}$ (just after step 2). The initial preconditioned control vector is filled with zeros and corresponds to $\mathbf{x_0} = \mathbf{x_b}$.

3.2 Observation operator and transport model

The observation operator (H in Equation 2) groups the ensemble of operations to compute the CO_2 concentrations corresponding to a given control vector. In our case, this covers the disaggregation from the monthly fluxes in the control vector to a 3-hourly temporal resolution, the addition of prescribed fluxes (fossil, ocean and biomass burning categories), their transport to the observations location and the addition of background concentrations.

In this first implementation of CO_2 inversions with LUMIA, we opted for a regional transport model based on pre-computed observational response functions (footprints):

$$y_m^i = y_{bg}^i + \sum_i \sum_c \langle \mathbf{K_i}, \mathbf{f_c} \rangle \tag{5}$$

where the footprint $\mathbf{K_i}$ stores the sensitivity of the observation y^i to the surface fluxes $\mathbf{f_c}$ (with the index c referring to the flux category), and y_{bg}^i is the contribution of background and historical fluxes to the model estimate y_m^i of observation y^i.

The flux vectors $\mathbf{f_c}$ are constructed from the control vector for the optimized flux category (NEE) and prescribed for the other fluxes. The relation between the control vector and the $\mathbf{f}_{nee}$ is given by:

$$\begin{aligned} \mathbf{f}_{nee}^t &= \frac{\mathbf{x}^m - \mathbf{x}_b^m}{n_t} + \mathbf{f}_{0_{nee}}^t \\ \mathbf{x}_b^m &= \sum_t \mathbf{f}_{0_{nee}}^t \end{aligned} \tag{6}$$

where $\mathbf{f}_{nee}^t$ is the NEE flux map at time step t of the month m, $\mathbf{f}_{0_{nee}}^t$ is the corresponding prior NEE map, $\mathbf{x}^m$ and $\mathbf{x}_b^m$ are the control vector and prior control vector components corresponding to month m at the same spatial coordinates, and n_t is the number of three-hourly intervals in the month m. In other words, the inversion adjusts an offset to the prior, high temporal resolution fluxes.

The adjoint operations corresponding to Equations 5 and 6 are summarized by

$$\mathbf{x}_{adj}^m = \frac{1}{n_t} \sum_t^{n_t} \sum_i \mathbf{K}_i^t \delta y^i \tag{7}$$

with δ_y^i the model-data mismatches weighted by their uncertainties (See Section 3.3.1).

Since $\mathbf{K}$ and y_{bg} are constant throughout the inversion iterations, they can be pre-computed, which reduces the transport computations to a set of very simple matrix operations. This tremendously reduces the computational cost of the inversions but

increases the I/O and storage requirements (one response function $\mathbf{K}$ must be stored for each observation and is read at each forward and adjoint iteration).

Although the operations described above are all part of the observation operator ('lumia.obsoperator' module), the actual forward and adjoint flux transport (Equation 5 and its adjoint counterpart $f^{adj} = H^{adj}(dy)$) are performed by the transport model ('transport' python library), called as a subprocess. The transport model relies on pre-computed observation footprints (response functions) and time series of background concentrations:

3.2.1 Response functions (regional transport model)

The response functions ($\mathbf{K}$) were computed using the FLEXPART 10.0 Lagrangian transport model (Seibert and Frank, 2004; Stohl et al., 2010). FLEXPART simulates the dispersion, backwards in time from the observation location, of a large number of virtual air "particles". The response function $\mathbf{K}_i^{\phi}$ corresponds to the aggregated residence time of the particles released for observation y^i, in a given space-time grid box ϕ of the regional inversion, and below a threshold altitude layer arbitrarily set to 100m).

The simulations were driven by ECMWF ERA-Interim reanalysis, extracted at a 3-hourly temporal resolution, and on a $0.5° \times 0.5°$ horizontal resolution, on a regional domain ranging from $25°$W, $23°$N to $45°$W, $83°$N, slightly larger than the inversion grid, which allows for some accounting of particles re-entry (i.e. when an air mass leaves the inversion domain, and re-enters it later, which is not accounted for in the background).

One set of 3-hourly response functions was computed for each observation, up to seven days backward in time (less if all the particles leave the domain sooner). For plain or low altitude sites (see Table 1), the particles were released from the sampling height above ground of the observations. For high altitude sites (around which the orography is unlikely to be correctly accounted for), the particles were released from the altitude above sea level of the observation sites.

The response functions are stored in HDF5 files, following a format described in SI.

3.2.2 Background concentrations (global transport model)

The background CO_2 concentrations (y^{bg} in Equation 5) result from the transport of CO_2-loaded air masses from outside the regional inversion domain to the observation sites. One approach to compute these background concentrations has been proposed by Rödenbeck et al. (2009), and consists in extracting background concentrations time series at the observation sites from the model output of a global, coarse resolution Eulerian transport model, driven by a set of inversion-derived CO_2 fluxes.

The background extraction is done in three steps: - A global, coarse resolution inversion is performed, constrained by a realistic set of prior CO_2 fluxes f_{apri}^{glo}, an initial atmospheric distribution of CO_2 concentrations (C^{ini}), a set of global, background surface CO_2 observations and a subset of the observations to be used later in the regional, high resolution CO_2 inversion. The aim of this step is to obtain a set of CO_2 fluxes f^{glo} that leads to a very realistic atmospheric CO_2 distribution in and around the regional inversion domain (RID). The accuracy of the fluxes themselves has less importance. - The CO_2 concentrations y^{tot} corresponding to the transport of the optimized coarse resolution fluxes f^{glo} to the observation sites within the RID are computed using a forward run of the global transport model used in step 1. The foreground CO_2 concentrations

y^{fg} are computed using a modified version of that same model, in which the fluxes and concentrations are maintained as zero at all times outside the regional domain, so that the concentrations y^{fg} result only from the transport of the fraction of the fluxes f^{glo} that is within the RID. - The background CO_2 concentrations are obtained by subtraction of the foreground concentrations to the total ones ($y^{bg} = y^{tot} - y^{fg}$).

The underlying assumptions is, that, by the time the air masses originating from outside the RID reach the observation sites, existing high resolution patterns of CO_2 at the regional domain boundaries concentrations would have been dispersed, and therefore the field of background CO_2 concentrations within the RID can be well represented with a coarse resolution transport model. On the other hand, this background CO_2 distribution should be as realistic as possible (within the limits of the model resolution), especially in and around the boundaries of the foreground domain, therefore the use of an inversion in step 1 above. We refer to Rödenbeck et al. (2009) for a much more complete description of the approach.

We implemented the Rödenbeck et al. (2009) approach in a TM5 model setup (Huijnen et al., 2010) with the initial global inversion (step 1) performed in a TM5-4DVAR setup, based on (Basu et al., 2013). The NEE flux is optimized monthly on a global $6^\circ \times 4^\circ$ grid, and three additional prescribed CO_2 flux categories are accounted for (fossil fuel, biomass burning and ocean sink). It covers the entire period of the LUMIA inversion, plus six extra months at the beginning and one at the end to limit the influence of the initial condition and to ensure that the background concentrations in the last month of the LUMIA inversion are well constrained by the observations (observations provide important constraints on the fluxes from the preceding month).

The inversion is constrained by flask observations from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network (Dlugokencky et al., 2019) outside the European domain, and by a subset of the observations used for the regional inversion within the European domain (see Section 3.3 for references, and Table SI1 for a full list of the sites used in that step).

Since the focus of this inversion is to produce a realistic CO_2 distribution around the European domain, the choice of a prior matters a lot less than the selection of observations. For practical reasons, prior fluxes from the CarbonTracker 2016 release were used (Peters et al., 2007): the NEE prior is generated by the SibCASA model (Schaefer et al., 2008); fossil fuel emissions spatially distributed according to the EDGAR4.2 inventory (https://edgar.jrc.ec.europa.eu/overview.php?v=42); biomass burning emissions are based on the GFED4.1s product (Van Der Werf et al., 2017) and the ocean flux is based on the Takahashi et al. (2009) climatology. We refer to the official CarbonTracker 2016 documentation (https://www.esrl.noaa.gov/gmd/ccgg/carbontracker/CT2016) and to references therein for further documentation on these priors.

The total (y^{tot}) and foreground (y^{fg}) CO_2 concentration time series at the observation sites are extracted using a modified forward TM5 run implementing the step 2 of the background extraction approach described above. The foreground and total CO_2 time series were saved for each observation site, both as continuous (every 30 minutes) concentration time series, sampled at the actual altitude (above sea level) of the observation site, but also as vertical profiles between the surface (as defined in the TM5 orography) and 5000 m.a.s.l (with a vertical resolution of 250 m and a temporal resolution of 30 minutes). The latter is used to construct a part of the observation uncertainties.

3.3 Observations and observational uncertainties

Observations from the GLOBALVIEWplus 4.2 obspack product were used in the inversions (NOAA Carbon Cycle Group ObsPack Team, 2019). For the year 2011, the product includes observations from 26 sites within our regional domain (in addition to observations from mobile platforms, which were not used). Continuous observations are available at 18 of these 26 sites and nine sites are high altitude. Most of these observation sites are now part of the European ICOS network. A list of the sites (coordinates, observation frequency, sampling height and data provider) is provided in Table 1, and the location of the observations is also reported in Figure 2.

3.3.1 Observation uncertainties

The observation uncertainty matrix ($\mathbf{R}$) accounts for both the measurement uncertainties (ε_{obs}) and the model representation uncertainty (ε_H, i.e. the incapacity of the model to represent perfectly well the observations, even given perfect fluxes). In theory, the diagonal of the matrix stores the absolute total uncertainty associated to each observation while the off-diagonals should store the observation error correlations. In practice, these correlations are difficult to quantify, and the size of the matrix would anyway make it impractical to invert. The off-diagonals are therefore ignored in our system (as in most similar inversion setups) and the observation uncertainty is stored in a simpler observation error vector, ε_y.

Our inversion system uses an observation operator that decomposes the background and foreground components of the CO_2 mixing ratio, therefore the model uncertainty can itself be decomposed in foreground and background uncertainties:

$$\varepsilon_y = \sqrt{\max\{\varepsilon_{obs}, \varepsilon_{obs}^{min}\}^2 + \varepsilon_{bg}^2 + \varepsilon_{fg}^2} \tag{8}$$

The instrumental error (ε_y) is provided by the data providers for most of the observations, and typically ranges between 0.1-0.7 ppm (see Figure 3). We enforced a minimum instrumental error (ε_{obs}^{min}) of 0.3 ppm for all the observations.

The model representation error can not be formally quantified, as this would require knowing precisely the CO_2 fluxes that the inversion is attempting to estimate. One can, however, assign representation error estimates (foreground and background) based, in particular, on assumptions of situations that would normally lead to a degradation of the model performances (for instance late-night/early-morning observations, with a development of the boundary layer that may not be well captured by the model, or observations in regions with a complex orography). Transport model comparisons can also provide representation error estimates based on the difference in their results.

3.3.2 Foreground model uncertainties

As described in Section 3.2.2, the TM5 simulation used for computing background CO_2 concentrations, also computes the foreground concentrations at each observation site. We performed a forward transport simulation with the regional transport model in LUMIA, using both the background concentrations and the foreground fluxes from that TM5 simulation, so that the

two simulations differ only by their regional transport model. A comparison between the concentrations computed by the two models is shown in Figure 4. The bias between the two models is very contained during the summer months (it is below 0.2 ppm from April to September, and goes as low as 0.01 ppm in July), but rises during the winter months (up to 1.45 ppm in November). The mean average difference between the two simulations is also much larger in winter: it ranges from 0.82 ppm in September to 4.3 ppm in November, with a yearly average of 3.3 ppm.

This comparison is not a formal performance assessment of either TM5 or of the FLEXPART-based transport used in LUMIA, and it particular the bias should be interpreted with care as the sign of the total net foreground flux changes during the year (which mechanically leads to a change of the sign of the bias). Nonetheless, it provides an indication on the order of magnitude of the foreground model transport errors. We use the absolute differences between the two models as a proxy for ε_{fg}.

3.3.3 Background model uncertainties

Background concentrations are expected to be accurately estimated by the global TM5 inversion when the dominant winds are from the West and that any signal from a strong point CO_2 source or sink has had time to dissipate along the air mass trajectory over the Atlantic Ocean. In less favourable conditions, there can be entries of less well-mixed air inside the domain, in particular in case of Easterly winds or in events of re-entry of continental air that would have previously left the domain. These events are less likely to be well captured by the TM5 inversion and should be attributed a higher uncertainty.

There is no perfect and easy way to detect these events, but one of their consequences would be a less homogeneous background CO_2 distribution around the observation sites when they occur. As part of the TM5 simulation, vertical profiles of background concentrations were stored for each observation (from the surface to 5000 m.a.s.l, at a 250m vertical resolution). We set the background uncertainty of each observation (ε_{bg}) to the standard deviation of its corresponding background CO_2 vertical profile. ε_{bg} is on average 0.36 ppm, one order of magnitude lower than ε_{fg}, and it is also more constant (it ranges between 0.01 and 3.6 ppm). Note that these statistics are computed before the observation selection procedure, described in the following section. The different components of the observation uncertainty are compared in Figure 3 for two representative sites.

3.3.4 Observation selection

The inversions are performed on a subset of the observations included in the obspack product. Only observations for which the transport model simulation is expected to result in accurate concentrations are kept. In practice, one of the main difficulties of transport modelling is to correctly compute the mixing of air in the lower troposphere below the boundary layer. The lowest model representation error is expected for observations that are either within the boundary layer when it is most developed (in the afternoon), or well above the boundary layer for high-altitude sites (during the night). For each site with continuous observations, we selected only observations sampled during the time range for which the model is expected to perform the best. The time ranges are based on the "`dataset_time_window_utc`" flag in the metadata of the observation files from the obspack. For sites with discrete sampling, all observations were used.

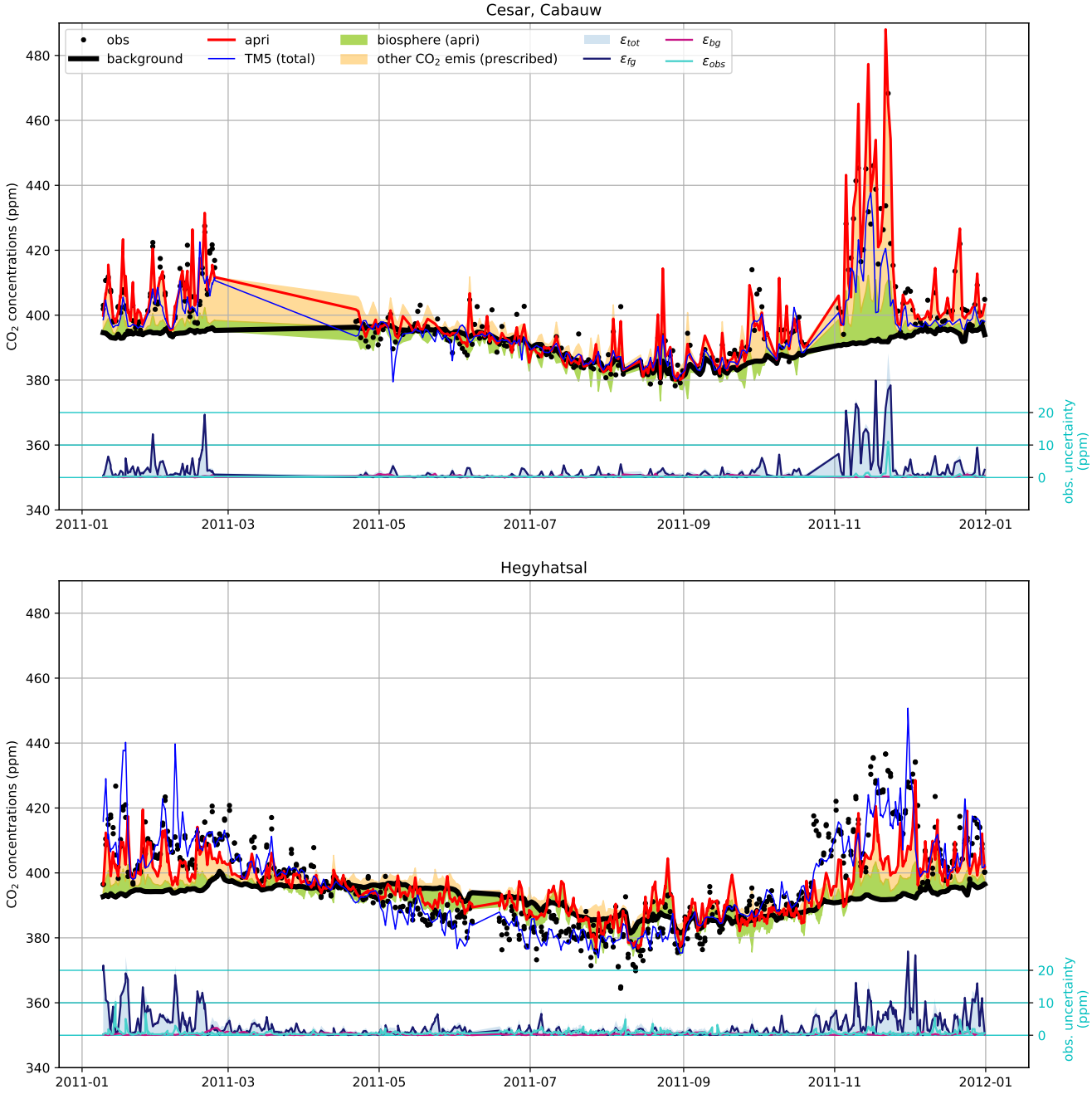

Figure 3. Decomposition of the modelled mixing ratio and of the observation uncertainties at two sites (Cabauw, The Netherlands and Hegyhatsal, Hungary). The "TM5 total" line is the concentration computed in the coarse resolution TM5 inversion from which the background (thick black line) is extracted. The LUMIA prior concentration is shown in red and the green and orange shaded areas show respectively the contribution of the prior biosphere flux and of the other CO_2 fluxes to the difference between that prior and the background. The lower series of lines in each plot (with y-axis on the right) shows the total observation uncertainty (blue shaded area), and the contributions of the foreground, background and observational uncertainties.

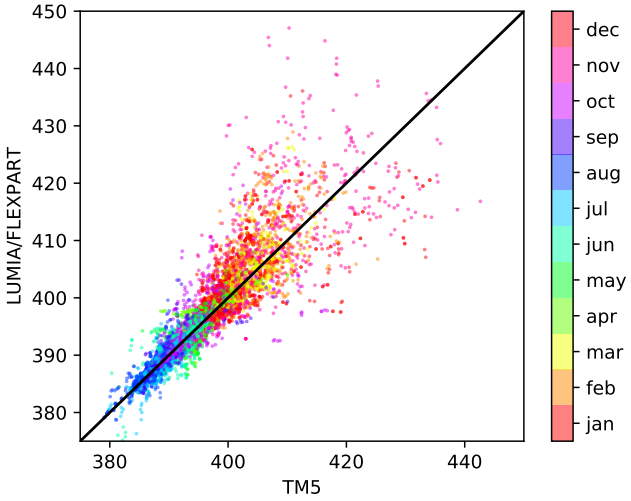

Figure 4. LUMIA (CO_2 concentrations obtained with TM5-FLEXPART vs. with TM5, using the CO_2 fluxes used as prior of the TM5 inversion (background). The color of the dots show the observation month.

A second filter, used in some of the inversions and in the background TM5 inversion (Section 3.2.2), is the limitation of a maximum of one observation per 24 hours at each observation site (the one that has the lowest observation uncertainty, as per the definition in the previous above). This is justified by the fact that two observations at a same site, within a small time interval have strongly correlated model representation errors and do therefore not provide independent information. In the absence of a proper accounting of observational error correlations, it may be preferable to limit the number of assimilated observations. The use of this second observation filter is discussed in further details in Section 5.

3.4 Prior and prescribed fluxes

In addition of the Net Ecosystem Exchange (NEE, net atmosphere-land CO_2 flux) that is optimized in the inversions, the simulations also account for anthropogenic CO_2 emissions (combustion of fossil fuels, bio fuels and cement production), for biomass burning emissions (large scale forest fires) and for the ocean-atmosphere CO_2 exchanges.

The NEE prior is taken from simulations of the LPJ-GUESS and ORCHIDEE vegetation models: in the OSSEs (Section 4) ORCHIDEE fluxes are used as prior and LPJ-GUESS fluxes are used as truth, while in inversions against real data LPJ-GUESS fluxes are used as prior. Both vegetation models provide 3-hourly fluxes, on a horizontal $0.5° \times 0.5°$ grid.

LPJ-GUESS (Smith et al., 2014) is a dynamic global vegetation model (DGVM), which combines process-based descriptions of terrestrial ecosystem structure (vegetation composition, biomass and height) and function (energy absorption, carbon and nitrogen cycling). The vegetation is simulated as a series of replicate patches, in which individuals of each simulated plant functional type (or species) compete for the available resources of light and water, as prescribed by the climate data. The model

Code	Name	Lat ($^\circ E$)	Lon ($^\circ N$)	Alt (m.a.s.l)	Intake Height (m.a.g.l)	Nobs	Time range (h)	sets	Data provider
BAL	Baltic Sea	55.35	17.22	3	25	53	all	P	1
BSC	Black Sea, Constanta	44.18	28.66	0	5	17	all	P	1
CES200	Cesar, Cabauw	51.97	4.93	-1	200	306	11-15	P	2
CIB005	Centro de Investigacion de la Baja Atmosfera (CIBA)	41.81	-4.93	845	5	49	*	A	1
CMN	Mt. Cimone Station	44.18	10.70	2165	12	549	23-3	A	3
CRP	Carnsore Point	52.18	-6.37	9	14	589	12-16	P	4
ELL	Estany Llong	42.57	0.95	2002	3	8	11-15	A	5
HEI	Heidelberg	49.42	8.67	116	30	632	11-15	P	6
HPB054	Hohenpeissenberg	47.80	11.02	936	54	47	all	A	1
HUN115	Hegyhatsal	46.95	16.65	248	115	685	11-15	P	7
JFJ	Jungfraujoch	46.55	7.99	3570	10	461	23-3	A	8
KAS	Kasprowy Wierch	49.23	19.98	1989	5	481	23-3	A	9
LMP005	Lampedusa	35.52	12.62	45	5	35	all	P	1
LMP008	Lampedusa	35.52	12.62	45	8	418	10-14	PA	10
LUT	Lutjewad	53.40	6.35	1	60	289	11-15	P	11
MHD024	Mace Head	53.33	-9.90	5	24	352	12-16	P	12
OPE120	Observatoire Perenne de l'Environnement	48.56	5.50	390	120	405	11-15	P	12
OXK163	Ochsenkopf	50.03	11.81	1022	163	48	all	A	1
PAL	Pallas-Sammaltunturi	67.97	24.12	565	5	654	22-2	PA	13
PRS	Plateau Rosa Station	45.93	7.70	3480	10	445	23-3	A	14
PUI	Puijo	62.91	27.65	232	84	170	11-15	P	13
PUY010	Puy de Dome	45.77	2.97	1465	10	409	23-3	A	12
PUY015	Puy de Dome	45.77	2.97	1465	15	141	23-3	A	12
SSL	Schauinsland	47.92	7.92	1205	12	625	23-3	A	15
TRN180	Trainou	47.96	2.11	131	180	539	11-15	P	12
TTA	Tall Tower Angus	56.56	-2.99	400	222	435	12-16	PA	16
WAO	Weybourne, Norfolk	52.95	1.12	20	10	1078	12-16	P	17
WES	WES	54.93	8.32	12	0	1377	11-15	P	18

Table 1. Observation sites used in the inversions. Data providers: 1:NOAA Carbon Cycle Group ObsPack Team (2019); 2:Vermeulen et al. (2011); 3:Ciattaglia et al. (1987); 4:D. Dodd (EPA Ireland); 5:J.A. Morgui and R. Curcoll (ICTA-UAB, Spain); 6:Hammer et al. (2008); 7:Haszpra et al. (2001); 8:Uglietti et al. (2011); 9:Rozanski et al. (2014); 10:A. G. di Sarra (ENEA, Italy); 11:van der Laan et al. (2009); 12:Yver et al. (2011); 13:Hatakka et al. (2003); 14:F. Apadula (RSE, Italy); 15:Schmidt (2003); 16:Ganesan et al. (2015); 17:Wilson (2012); 18:K. Uhse (UBA, Germany)

is forced using the WFDEI meteorological data set (Weedon et al., 2014) and produces 3-hourly output of gross and net carbon fluxes.

ORCHIDEE is a global processed-based terrestrial biosphere model (initially described in Krinner et al. (2005)) that computes carbon, water and energy fluxes between the land surface and the atmosphere and within the soil-plant continuum. The model computes the Gross Primary Productivity with the assimilation of carbon based on the Farquhar et al. (1980) for C3 plants and thus account for the response of vegetation growth to increasing atmospheric CO_2 levels and to climate variability.

Category	Product	Original resolution	Data provider	Total (min/max) flux (PgC/year)
Biosphere	LPJ-GUESS	$0.5° \times 0.5°$; 3-hourly	Lund University	-0.33 (-2.65 / 1.85)
Biosphere	ORCHIDEE	$0.5° \times 0.5°$; 3-hourly	LSCE (P. Peylin, pers. comm)	-0.28 (-3.73 / 2.14)
Fossil	EDGARv4.3	$0.1° \times 0.1°$; hourly	ICOS-CP + JRC	1.53
Ocean	CarboScopev (oc_v1.7)	$5° \times 3.83°$; daily	Rödenbeck et al. (2013)	-0.11 (-0.05 / 0.01)
Fires	GFEDv4	$0.5° \times 0.5°$; 3-hourly	Van Der Werf et al. (2017)	0.01

Table 2. Prior and prescribed CO_2 fluxes. Min/Max values are provided for the fluxes that have both positive and negative components, and correspond to the minimum and maximum values of the 3-hourly flux aggregated over the entire domain, in PgC/year.

The land cover change (including deforestation, regrowth and cropland dynamic) were prescribed using annual land cover maps derived from the Harmonized land use data set (Hurtt et al., 2011) combined with the the ESA-CCI land cover products. The net and gross CO_2 fluxes used for this project correspond to the one provided for Global Carbon Project inter-comparison (Le Quéré et al., 2018) with a model version that was updated recently (Peylin et al., in preparation).

Fossil fuel emissions are based on a pre-release of the EDGARv4.3 inventory for the base year 2010 (Janssens-Maenhout et al., 2019). This specific dataset includes additional information on the fuel mix per emission sector and thus allows for a temporal scaling of the gridded annual emissions for the inversion year (2011) according to year-to-year changes of fuel consumption data at national level (bp2, 2016), following the approach of Steinbach et al. (2011). A further temporal disaggregation into hourly emissions is based on specific temporal factors (seasonal, weekly, and daily cycles) for different emission sectors (Denier van der Gon et al., 2011).

The ocean-atmosphere flux is taken from the Jena CarboScope v1.5 product, which provides temporally and spatially resolved estimates of the global sea-air CO_2 flux, estimated by fitting a simple data-driven diagnostic model of ocean mixed-layer biogeochemistry to surface-ocean CO_2 partial pressure data from the SOCAT v1.5 database (Rödenbeck et al., 2013).

A biomass burning flux category was also included in the inversion, based on fluxes from the Global Fire Emission Database v4 (Giglio et al., 2013). In our European domain biomass burning emissions are negligible regarding the other CO_2 emission sources, however, we include it for completeness.

All fluxes are regridded on the same $0.5° \times 0.5°$, 3-hourly resolution (by simple aggregation or re-binning). A summary of the prior fluxes sources, original resolution and yearly totals is provided in Table 2.

3.4.1 Prior uncertainties

The background error covariance matrix ($\mathbf{B}$ in Equation 2) is constructed following the "correlation length" approach used in many other inversion systems (e.g. Houweling et al. (2014); Thompson et al. (2015); Chevallier et al. (2005)): The error covariance between fluxes $\mathbf{x_1}$ and $\mathbf{x_2}$ at grid cells with coordinates $p_1 = (i1, j1, t1)$ and $p_2 = (i2, j2, t2)$ is defined as:

$$cov(\mathbf{x_1}, \mathbf{x_2}) = \sigma^2_{\mathbf{x_1}} . \sigma^2_{\mathbf{x_2}} e^{-(d(p_1,p_2)/L_h)^2} e^{-|t_1-t_2|/L_t} \quad (9)$$

where $\sigma^2_{\mathbf{x_1}}$ and $\sigma^2_{\mathbf{x_2}}$ are the variances assigned to the prior monthly NEE at coordinates p_1 and p_2, and L_h and L_t are correlation lengths, which define how rapidly the correlation between two components drops as a function of their distance in time and space.

The true uncertainty of the prior fluxes ($\sigma^2_{\mathbf{x}}$) is difficult to evaluate and is therefore constructed on reasonable but arbitrary assumptions. We tested several approaches, further discussed in Section 4: 1) Scaling the uncertainties linearly to the absolute next monthly flux; 2) Scaling the uncertainties to the absolute net 3-hourly flux (and then cumulating these 3-hourly uncertainties to the monthly scale); 3) Enforcing constant monthly uncertainties throughout the year, at the domain-scale.

3.5 Inversions performed

We performed two ensembles of inversions. The first ensemble consists of Observing System Simulation Experiments (OSSEs) to assess the theoretical performance of the system. Here, the LPJ-GUESS NEE dataset was taken as an arbitrary truth, and a dataset of synthetic pseudo-observations was generated at times and locations of the actual observations listed in Table 1, by forward propagation of the "true" NEE flux with the transport model (including also the contributions of non-optimized fluxes listed in Section 3.4). Random perturbations were then added, to mimic the measurement error ($y = y^{truth} + \mathcal{N}(0, \sigma^2_y)$, with σ^2_y the uncertainty of each observation as defined in the matrix $\mathbf{R}$).

The OSSEs use this set of pseudo-observations as observational constraint and the ORCHIDEE NEE dataset as a prior. The reference OSSE, SRef, uses a prior error covariance matrix ($\mathbf{B}$) constructed with prior uncertainties set to 25% of the absolute prior value ($\sigma^2_{\mathbf{x_b}} = 0.25|\mathbf{x_b}|$) and with covariances constructed from a horizontal correlation length (L_H) of 200 km and a temporal correlation length (L_t) of 30 days. In the sensitivity tests we vary the correlation lengths (SC.100 and SC.500), the prescribed prior uncertainties (SE.3H, SE.3Hcst, SE.x2) and the extent of the observation network (SO.A, SO.P).

The second ensemble is essentially identical to the OSSE ensemble of inversions, except that it is using real observations, and the LPJ-GUESS flux dataset as a prior. The details of the two ensembles of inversions are listed in Table 3

4 OSSEs

We first analyze the capacity of SRef to reconstruct various characteristics of the "true" LPJ-GUESS NEE fluxes (monthly and annual NEE budget, aggregated at spatial scales ranging from the entire domain down to single pixels). Then we use a series of sensitivity experiments to verify how sensitive the results are to a range of reasonable assumptions in the inversion settings.

4.1 Reference inversion (SRef)

Figure 5 shows monthly and annual time series of NEE and NEE error (with respect to the prescribed truth) aggregated over the entire domain.

At the domain-scale, the prior estimate for the annual NEE is very close to the "truth" (respectively -0.28 and -0.34 PgC/year), but the amplitude in the prior is more than double that of the truth, with monthly NEE ranging from +0.26 PgC/month in October to -0.66 PgC/month in June in the prior, compared to +0.08 PgC/month (in January) to -0.29 PgC/month

Simulation	Prior	Observations	$\sigma^2_{\mathbf{x}}$	L_h	L_t
SRef	ORCHIDEE	pseudo	25% of monthly prior	200 km	1 month
SC.500	-	-	-	500 km (exp)	-
SC.100	-	-	-	100 km (exp)	-
SE.x2	-	-	50% of monthly prior	200 km	-
SE.3H	-	-	12.5% of 3-hourly prior	200 km	-
SE.3Hcst	-	-	12.5% of 3-hourly prior, scaled to the same total value every month	-	-
SO.P	-	pseudo (set P)	-	200 km	-
SO.A	-	pseudo (set A)	-	200 km	-
RRef	LPJ-GUESS	real	25% of monthly prior	200 km	1 month
RC.500	-	-	-	500 km (exp)	-
RC.100	-	-	-	100 km (exp)	-
RE.x2	-	-	50% of monthly prior	200 km	-
RE.3H	-	-	12.5% of 3-hourly prior	200 km	-
RE.3Hcst	-	-	12.5% of 3-hourly prior, scaled to the same total value every month	-	-
RO.P	-	real (set P)	-	200 km	-
RO.A	-	real (set A)	-	200 km	-

Table 3. List of inversion experiments performed. The R and B letters in the Observations column refer respectively to a random error perturbation (R, proportional to the assigned individual uncertainty of each observation) and to a systematic bias (B), described in the main text of Section 4. The restricted observation sets A and P are reported in Table 1.

(in May) in the truth. In total, the absolute prior error slighly exceeds 3 PgC and peaks in June and July and is the lowest in December-February.

The inversion improves the estimation of the seasonal cycle at the domain scale, with a seasonal cycle amplitude reduced to a range of -0.36 PgC (May) to +0.16 PgC (December), much closer to the truth and the absolute error is reduced by nearly 40% to 1.87 PgC/year. However since the positive flux corrections in the summer months largely exceed the negative corrections from September to April, this results in a strong degradation of the annual European NEE estimate, with a near-balanced posterior flux of -0.05 PgC/year. Figure 6 illustrates the spatial distribution of the error reduction. While the largest prior errors are found north of the Black Sea and in North-Africa, the error reduction is rather homogeneous, except for North-Africa and Turkey (which are not really constrained by the observation network), and some patches in Western Europe (mainly in the UK, but also in Ireland, France and the Benelux) where the error actually increases. In total, these localized error enhancements amount to 0.16 PgC (lower panel of Figure 5). These isolated occurrences of error enhancements are not a sign of malfunction of the inversion system, but they highlight its limitations: they result from attributions of flux corrections to the incorrect grid cells, which can happen if the resolution of the inversion is not adapted to the constraints provided by the observation network (i.e. smoothing and aggregation errors, as defined in Turner and Jacob (2015)).

Although our control vector contains the flux estimates at the native spatial resolution of the transport model, the effective resolution of the inversion is further constrained by the covariances contained in the prior error-covariance matrix $\mathbf{B}$. Furthermore, the fluxes are only optimized monthly, while the actual prior error varies at a 3-hourly resolution. It may there-

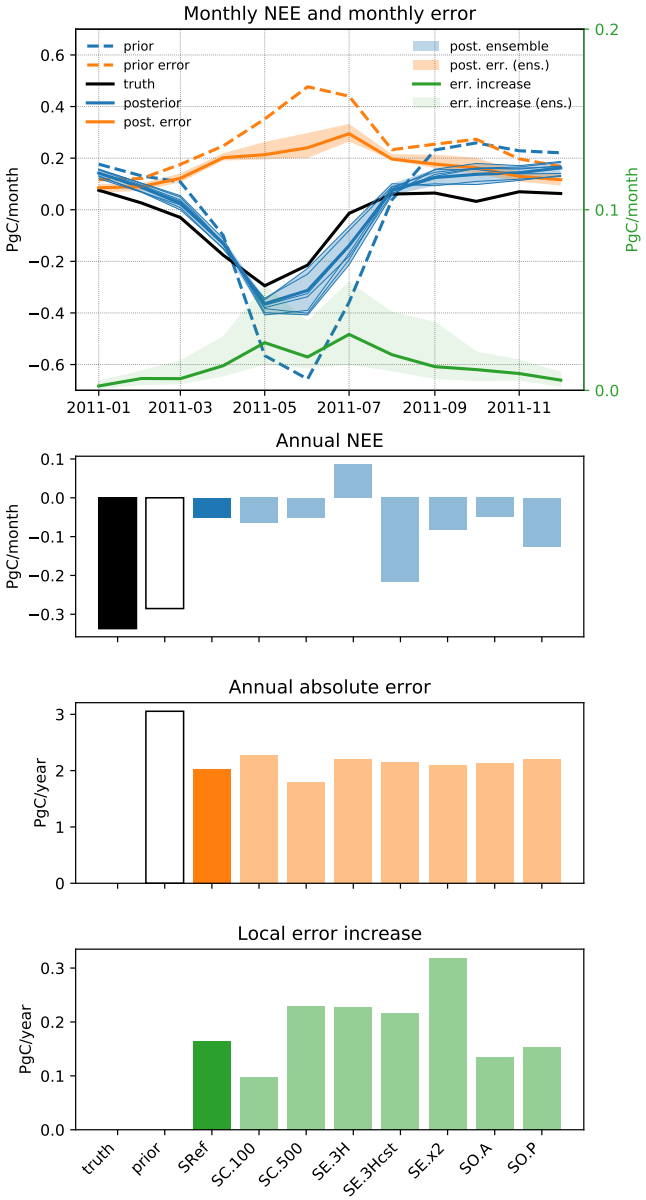

Figure 5. Upper row, left axis: Monthly prior NEE (dashed blue line), true NEE (solid black line), posterior NEE (blue), absolute prior error (dashed orange line) and posterior error (orange) ; Upper row, right axis: Total error increase (i.e. positive component of the error reduction, green). The SRef inversion is shown as solid lines, the ensemble is shown as a shaded area. Second, third and fourth rows: same variables, but aggregated annually.

fore be possible to reduce these errors by increasing the resolution of the inversion, but since the observation network is not homogeneous, this may lead to increased posterior error in other parts of the domain.

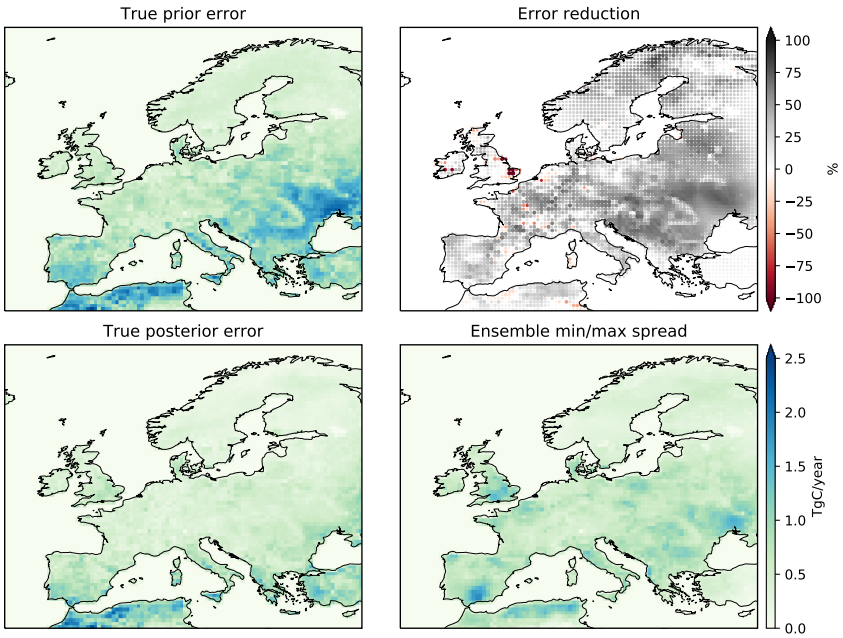

Figure 6. Left column: prior (top) and posterior (bottom) total error (with respect to the prescribed truth); Upper-right: percentage error reduction. The size of the dots is proportional to the size of the flux correction; Bottom-right: amplitude of the posterior ensemble spread.

4.2 Sensitivity tests

We performed a series of sensitivity tests to assess the robustness of the results. The sensitivity experiments can be grouped in three series: First, SC.100 and SC.500 test the sensitivity to the spatial covariance lengths used to construct $\mathbf{B}$. Then, SE.3H, SE.3Hcst and SE.x2 test the sensitivity of the results to the prior uncertainty themselves (i.e. diagonal elements of $\mathbf{B}$). Finally, SO.A and SO.P test the sensitivity of the results to the network density.

The total NEE flux, absolute error and error increase are shown in Figure 5, for the individual sensitivity experiments at the annual scale and as an ensemble shape for the monthly scale (the monthly-scale results of the individual simulations can be found in Figure SI2).

4.2.1 Sensitivity to the error distribution

Inversions SE.3H, SE.3Hcst and SE.x2 were designed to test the impact of the prescribed prior uncertainty vector (e.g. diagonal of $\mathbf{B}$) on the inversion:

- In SE.3H, the prior uncertainty is set proportional to the sum of the uncertainties on the 3-hourly fluxes: $\sigma_{\mathbf{x_b}} = \frac{0.13}{T} \sum_t^T |f_t|$. This avoids the situation where GPP and respiration are significant but compensate each other, leading to a near zero NEE as well as a near zero prior uncertainty, which can happen when the prior uncertainty is calculated following the approach used in SRef. The factor 0.13 was chosen to lead to a total annual uncertainty comparable to that of SRef. This leads to an overall redistribution of the uncertainties from the winter to the summer period, which is closer to the actual distribution of differences between the prior and truth fluxes (see Figure SI1).

- In SE.3Hcst, the prior uncertainty is computed as in SE.3H, but it is then scaled monthly, so as to lead to a flat distribution of the uncertainties across the year.

- In SE.x2, the prior uncertainty is simply doubled compared to SRef.

SE.3Hcst leads to an improved value of the annual budget of NEE at the domain scale, but this is due to a poorer estimation of the summer fluxes (since the uncertainty is lower in summer, the inversion sticks more to its prior). On the contrary, SE.3H leads to further degradation of the annual budget, without achieving better performances than SRef at the monthly scale. For both inversions, this translates into a slightly larger total posterior error (2.15 and 2.20 PgC/year, respectively, compared to 2.03 in SRef). The doubling of the prior uncertainty in SE.x2 allows it to depart more from the prior and to derive better domain-scale flux estimates, both monthly and annually, but it also leads to an increase in the "added error" (lower panel of Figure 5).

4.2.2 Sensitivity to the error covariance structure

Inversion SC.100 and SC.500 use prior error covariance matrices constructed using respectively shorter (100 km) and longer (500 km) horizontal correlation lengths (L_H) than SRef. The longer covariance length in SC.500 forces the inversion to favour large-scale, low amplitude flux corrections over localized strong adjustments. Since the prior error follows a relatively homogeneous pattern, SC.500 effectively produces a better estimation of the NEE, especially in Eastern Europe where the network is sparse (Figure SI4b). The opposite happens with SC.100, which tends to concentrate the flux adjustments in the vicinity of the observation sites.

At the domain scale, the annual budgets are nearly identical in SC.100, SC.500 and SRef. However the total error reduction is lower in SC.100 and higher in SC.500, compared to SRef (respectively 0.78, 1.28 and 1.02 PgC/year), but the "added error" is larger in SC.500 (0.23 PgC/year) and lower in SC.100 (0.10 PgC/year): this confirms the hypothesis that these are aggregation errors, that can be reduced by increasing the number of degrees of freedom in the inversion (for instance by reducing the covariance constraints).

4.2.3 Sensitivity to the observation network density

Compared to SRef, SO.A uses only high-altitude observations (plus LMP and TTA as these were the only sites available in their region) and SO.P uses only low altitude sites. In terms of annual budget, SO.P outperforms most of the other inversions, but as

for SE.3Hcst, this results from poorer flux corrections in summer rather than from a better overall reduction of the uncertainties. On the contrary SO.A leads to results very comparable to SRef at the domain scale, with a nearly identical seasonal cycle and net annual flux. The net error reduction remains however slightly better in SRef (see also Figure SI1 for the seasonal cycles of SO.A and SO.P).

4.3 Evolution of the fit to the observations

The comparison of the prior and posterior model fit to observations is a classical diagnostics of atmospheric inversions (Michalak et al., 2017). The inversion is expected to improve the overall fit to the observation ensemble, and a lack of statistical improvement would generally be a sign of a malfunctioning inversion algorithm. At a finer scale, analysis of the when and where the representation of the observations is most improved (or degraded), can provide useful insights on the performances of the inversion (adequacy of the definition of uncertainties) and on those of the underlying transport model.

In the right panel of Figure 7, we compare the statistical distribution of prior and posterior observation fit residuals for inversion Sref. The plot confirms that the inversion leads to an overall improvement of the representation of observations, albeit modest (prior bias (model-obs): 0.2 ppm; posterior bias: 0.05 ppm; prior RMSE: 4.9 ppm; posterior RMSE: 3.75 ppm). The left panel shows the RMSE reduction at each observation site (the size of the dots is proportional to the number of assimilated observations at each site, and the color shows the net RMSE reduction). At all sites the inversion leads to improvements in the fit, but those are generally much more modest in Western Europe, which can be explained by the (coincidental) good performance of the prior in that region (see Figure 6), but also by the strong sensitivity of these sites to background concentrations. Sites in the UK and, in particular, Ireland sample very little continental air, which leaves little margin for the inversion to improve the representation of their observations.

The center panel of Figure 7 compares the RMSE reduction of inversion SRef to that of the other OSSEs. The best performances are logically achieved by SE.x2, which can depart much more from its prior than the other inversions. On the other hand, SC.100 systematically underperforms the ensemble, which is coherent with its poorer flux error reduction. In general however, the reduction of misfits are very similar and are not good indicators for the quality of the optimized fluxes.

5 Inversions with real observations

The OSSEs presented above neglect several complications of real inversions, in particular transport model errors (the observations were generated using the same transport model as the one used in the inversions). While it is not within the scope of this paper to quantify precisely these errors, we nonetheless performed a series of inversions constrained by real observations, to assess to which extent the characteristics of the inversions results identified with the ensemble of OSSEs remain under a more realistic situation.

The ensemble of inversions used here is identical to the OSSEs ensemble, except that real observations are used and that the LPJ-GUESS flux is used as a prior (instead of ORCHIDEE in the OSSEs). The inversion settings are reported in Table 3.

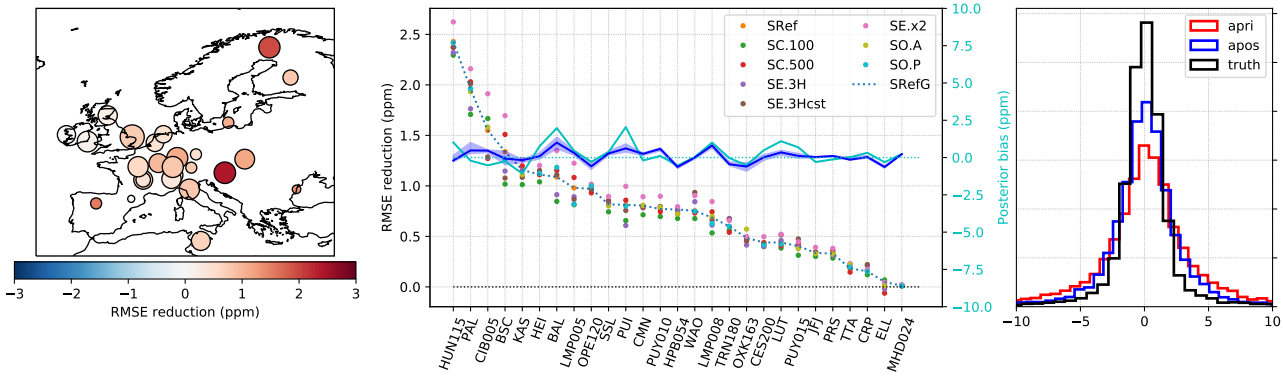

Figure 7. Left: Map of the observation sites in Sref, with the area of the dots proportional to the number of assimilated observation at each site, and the color proportional to the RMSE reduction (prior RMSE minus posterior RMSE). Center: RMSE reduction at each site, for the five sensitivity OSSEs. Right: distribution of observation residuals with the prior, posterior (SRef) and truth fluxes. The cyan and blue line plots in the center plot show the prior (blue) and posterior (cyan) mean biases at each site (right axis)

.

5.1 Posterior fluxes

The monthly and annual prior and posterior NEE are shown in Figure 8, for the reference RRef inversion and for the sensitivity tests. The inversion leads to a slight increase in the seasonal cycle amplitude, with a peak summer uptake increased by 24% in May (-0.36 PgC/month, instead of -0.24 PgC/month in the prior) and a nearly doubling of the CO_2 emissions in winter (+0.12 PgC/month instead of +0.07 PgC/month in the prior in November). It also leads to a delayed date for the change of sign of the net flux, both in the spring and in the autumn (the net prior flux becomes negative in March in the prior, and positive again in August, while it only becomes negative in April, and positive in October in RRef).

These monthly flux adjustments do not result in a change in the net annual flux (-0.33 PgC/year, both in the prior and in the RRef posterior). As seen when analysing the OSSEs results, the net annual budget is not well constrained by the inversions and the absence of change is here purely coincidental.

In contrast to the OSSEs, the transport model error is not zero, which may explain the slightly higher sensitivity of the results to the extent of the observation network: RO.P and RO.A differ by, on average, 0.02 PgC/month, the double of the average difference between SO.P and SO.A. However the overall spread of results in that second ensemble of inversions is on the same order of magnitude to that obtained with the OSSEs, with a monthly spread ranging from 0.02 PgC/month (January and September) to 0.07 PgC/month (March and August). This indicates that the conclusions of the OSSEs regarding the robustness of the results can be generalized to these inversions with real data.

Maps of the prior and posterior fluxes, as well as the flux adjustments obtained with RRef are shown in Figure 9, for three 4-months periods. The January to April and September to December periods correspond approximately to the time of the year when a positive NEE correction is obtained by the inversion, while May to August is the period when the inversion finds increased uptake compared to the prior. While at large scales, the inversion preserves well the spatial distribution of NEE,

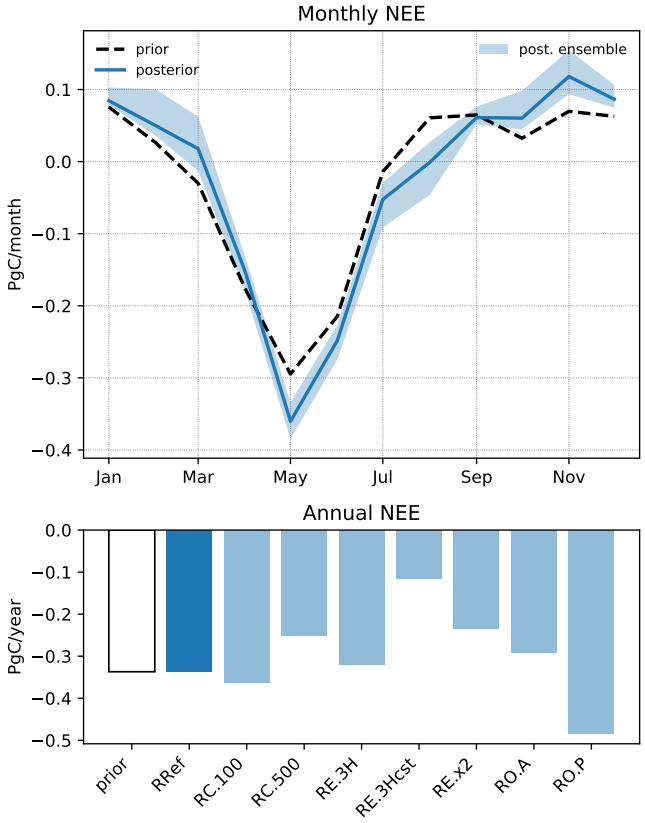

Figure 8. Top: Monthly prior (dashed black line) and posterior NEE (solid line: RRef; shaded blue area: ensemble spread); Bottom: Annual NEE for the prior LPJ-GUESS and the seven inversion posteriors.

the flux adjustment is not as homogeneous as what was obtained with the OSSEs (see also monthly flux adjustments maps in Figure SI5b).

505 The ensemble variability (lower row of Figure 9) is much higher than in the OSSEs in North-Western Europe (Northern France, Ireland and the UK), and in Hungary, around the Hegyhatsal observation site (see also Figure 6). In the latter case, this is mainly due to the inclusion or not of this site in the inversions (i.e. RO.A/RO.P inversions). The discrepancies in North-West Europe were already present in the OSSEs, but here with real observations the inversions additionally have to compensate for the inaccuracy of the transport model. In particular, errors in the prescribed background concentrations will have a stronger
510 impact on the optimized fluxes in the vicinity of sites that sample predominantly background concentrations, such as the sites in Ireland and the UK. But also, observation sites downwind of large urban areas are more susceptible to be impacted by errors in the prescribed fossil fuel emissions, either because the emission scenario itself is incorrect, or the transport model resolution is too coarse to correctly represent the impact of these emissions at the observation site.

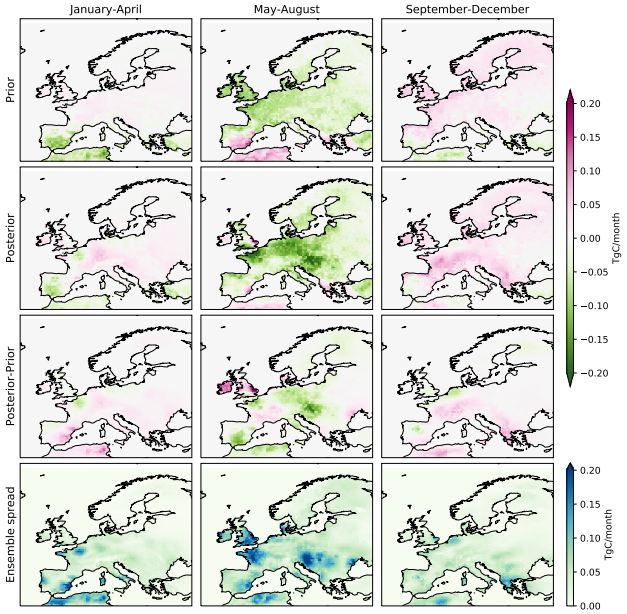

Figure 9. Total prior NEE (top), posterior NEE (second row) and NEE adjustment (third row) for the RRef inversion, and for three 4-months period (left to right); Lowe row: Posterior ensemble spread

5.2 Reduction of the observation misfits

Figure 10 provides an overview of the model-data mismatches for RRef and, at the site level, for the sensitivity experiments. As expected, the inversion leads to a reduction in the RMSE, from 5.8 ppm in the prior to 4.8 ppm in the posterior, and to a slight reduction of the mean bias (from -0.2 ppm to -0.1 ppm). These values are slightly larger than the ones obtained in the OSSEs, which is consistent with the presence of a non-perfect transport model and boundary conditions.

At the site level, the prior biases are more variable than in the OSSEs ensemble, from -9.1 ppm at Baltic Sea (BSC) to +2ppm at Ochsenkopf (OXK). The bias corrections remain very modest at most sites (the bias even slightly increase at a few sites). The large (7.5 ppm) bias at BSC (Black Sea) is computed from a very small number of observations is very small (17 in total, with observational errors up to 8 ppm), which have therefore very little weight in the inversion. The RMSE is generally reduced, except at ELL (Estany LLong, Spain) and OXK (Ochsenkopf, Germany), where the fit to the observations is slightly degraded. Both sites are located in relative proximity to other observation sites, with which their footprints overlap: the degradation of the RMSE results from contradictory constraints provided to the inversions by these different sites. The inversion does not have sufficient degrees of freedom to improve simultaneously the fit at all sites, and therefore degrades the fit to the OXK and ELL observations, which have only few observations (48 and 8, respectively). The problem is common to all the sensitivity runs, and the mean posterior biases are also very similar across the ensemble.

As seen with the OSSEs, a better performance in the fit to observations is not necessarily an indication of a more accurate optimized solution. The site-by-site analysis of the misfits might point to limitations of the transport operator, but a more in-depth analysis would be required, which is out of the scope of this paper.

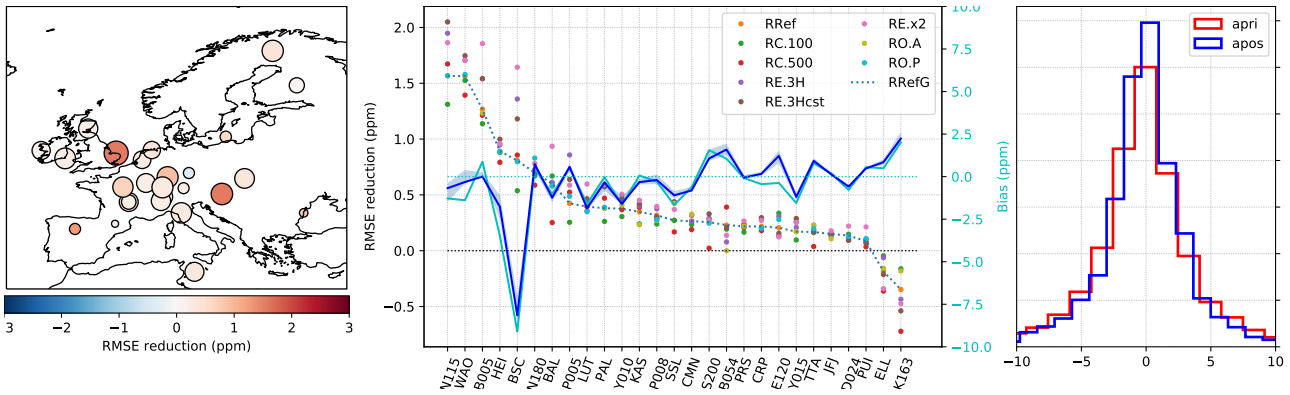

Figure 10. Left: Map of the observation sites in Rref, with the area of the dots proportional to the number of assimilated observation at each site, and the color proportional to the RMSE reduction (prior RMSE minus posterior RMSE). Center: RMSE reduction at each site, for the five sensitivity OSSEs. Right: Prior and posterior distribution of the observation mismatches in Rref (irrespective of the site)

.

6 Discussion and conclusions

We have setup an atmospheric inversion system based on an implementation of the variational inversion approach (Section 3.1) with a transport model based on an offline coupling between FLEXPART (high-resolution regional transport) and TM5 (coarse-resolution transport of the background fluxes and historical atmospheric CO_2 burden). The inversion was tested through a series of synthetic experiments and realistic inversions, which show that it is working as expected. In this section we discuss separately three aspects of the paper. First the inversion results themselves, then the TM5-FLEXPART coupling and finally the LUMIA system.

6.1 Inversion approach and results

We have setup a variational inversion framework, to optimize European NEE at a monthly, 0.5° scale. The setup is intentionally simple: the aim at this stage was to develop a robust technical base and to have a reference setup for future developments. The transport model is a transposition to TM5 and FLEXPART of the off-line coupling developed by Rödenbeck et al. (2009) for TM3 and STILT, and the optimization itself shares many similarities with existing inversion systems, e.g. TM5-4DVAR (Basu et al., 2013), TM3-STILT (Kountouris et al., 2018) or even PyVAR-CHIMERE (Broquet et al., 2011), which should facilitate the comparison of results with these systems.

The first inversion results suggest that the inversion system is working as expected. In the context of OSSEs, the inversions enable on average a 40% reduction of the flux error at the grid-cell, monthly scale, and the differences between the optimized

fluxes obtained from different sensitivity runs are in line with what could be expected from the different settings used. However, these local error reduction can be of opposite sign, and do not always add up to a net error reduction at larger scales. In particular, while the NEE estimate is generally always improved at the monthly scale, the positive corrections in summer are much stronger than the negative corrections in winter, which results in an overall degradation of the annual NEE.

Using an even month-to-month distribution of the uncertainties (SE.3Hcst inversion) leads to a more realistic annual estimate, but also to a higher occurrence of local degradations of the solution, which further complicates the interpretation of the results. This high sensitivity of the annual NEE to the different choices of prior uncertainty show that this metric is not well constrained in our inversions: unless additional constraints are introduced in the system, it will remain difficult to provide a more robust estimate of the annual European NEE than what can be obtained through other estimation techniques (bottom-up modelling, global inversions, etc.). In particular, making use of constraints on the large-scale gradients from the global inversion (background) to the regional one would be beneficial.

Another approach is to accept that this metric (the annual European NEE) can remain under-constrained in our inversions, and focus on the aspects of the solution that the inversions really improve: in the OSSEs, regardless of the specific inversion setup, the posterior provides a much more realistic depiction of the seasonal cycle of NEE, and of its spatial variability. Furthermore the results gain in consistency (i.e. become less sensitive to sensitivity experiments) where the observation network is dense, which is encouraging since the observation network in Europe has significantly expanded compared to the data selection used in this paper. The inversions provide relevant information that can help identifying specific shortcomings in the prior flux estimates, and the practice of performing an ensemble of sensitivity runs greatly helps in identifying the most robust features. It is here facilitated by the low marginal computational cost of computing additional inversions.

The OSSEs systematically lead to some degradation of the solution in the parts of the domain that are very densely covered by the observation network, which is counter-intuitive. It may be partly because the prior was already very close to the truth in this part of the domain, which makes it difficult for the inversion to further optimize the solution, but a complementary explanation is that the system may not have sufficient degrees of freedom to adjust the fluxes to simultaneously improve the fit at all observation sites. In particular, the optimization of monthly fluxes is very restrictive. The implementation of an optimization at a higher temporal resolution will therefore be an important next step. In addition, varying the resolution of the optimization according to the density of the observation network may also help (either by varying the resolution of the optimized fluxes, or by varying the covariance lengths in the prior error-covariance matrix).

The application of the same inversion approach to real observations leads to smaller flux adjustments than in the OSSEs. This could be a sign that the difference between the LPJ-GUESS prior (used in this second set of inversions) and the true fluxes is smaller than that between the prior and synthetic truth in the OSSEs, but the analysis of the observation misfits reduction also point to potential site-dependent transport model errors. One of the next steps towards improving our inversions will therefore have to be a thorough model calibration effort. In that sense, the flexibility of LUMIA with regards to the transport model is particularly adapted.

6.2 TM5-FLEXPART coupling

The inversions rely on an offline coupling between the FLEXPART Lagrangian transport model (for regional, high resolution transport) and TM5-4DVAR for providing background concentrations. The setup replicates the 2-step scheme of Rödenbeck et al. (2009) but with different models.

A succint comparison between this "TM5-FLEXPART" transport model and TM5 itself was performed (Section 3.3.1) and is used as a proxy for the transport model error. It doesn't show any global bias between the two models, but a possible seasonal offset towards the month of November. The prescribed observation uncertainties are scaled up to account for this possible larger model error, so the impact on inversions should be limited. Nonetheless, that possible seasonal bias would need to be investigated and accounted for before deriving scientific conclusions from inversions against real observations.

The choice of the models and of that specific coupling was driven in part by the perspective of exchanges with other groups using similar setups. In the current stage, replacing the FLEXPART response functions from another similar Lagrangian transport model (STILT(Lin et al., 2003), NAME (Jones et al., 2007), etc.) or the TM5 background time series by data generated with a different model (using either the same or a different technique to estimate background concentrations at the observation sites) is straightforward and will facilitate a better evaluation of the model performance.

Note also that the Rödenbeck et al. (2009) approach means that there is no 'hard' coupling between the two models meaning that there is no risk of having to use an older version of one model because of the lack of implementation of the coupling in newer code. This, of course, also facilitates the exchange of one transport model against another as mentioned above.

From a practical and technical point of view, the current setup presents the advantage of speed and scalability: the application of the transport operator is done independently for each observation and therefore can be distributed on as many CPUs as available. Inversions can thus be performed in very limited (user) time (5-8 hours wall time per inversion on 24 CPUs for the inversions in this paper). This time efficiency is critical for running not only single inversions, but inversion ensembles, which provide a better representation of the real uncertainties. Evolutions of the code for very large ensemble of observations (such as from satellite retrievals) may, however, benefit from further developments (aggregation of observations and footprints; reduction of the number of grid points where possible; etc.).

6.3 The LUMIA framework: conclusions and future perspectives

We have developed the LUMIA inversion framework, and performed a first set of inversions with it. The framework is initially designed for the purpose of performing regional CO_2 inversions in Europe, however it is designed and developed as a flexible and adaptable inversion system, which enables the easy exchange of major components of the system, such as the transport model or the minimization algorithm, to isolate and study their impact on the inversion results. LUMIA is designed to be transport-model agnostic, i.e. it is not constructed on top of an existing model and it calls the transport model via a well-defined interface.

Technically, the inversion framework presented in this paper includes three major components: the lumia python library, which contains most of the actual inversion code in the form of independent modules; a transport component, which relies

on pre-computed observation footprints and background concentrations; the inversion scripts themselves, which use the lumia library and the transport model to implement the inversion experiment.

The lumia python library defines an ensemble of classes corresponding to basic elements of the inversion setup, e.g. control vector, gradient descent algorithm, transport model interface, observations database, etc. The library is distributed on a git server and is installable via the standard pip tool, which means it can be installed in one single command on a new computer. Although the library is developed and designed for the purpose of inversions, we have made particular effort to ensure the modularity of the code: the different modules can be imported independently and be used to construct new experiments. On the short-term, this design facilitates the re-use of the code in pre-/post-processing steps of the inversions, as well as during the analysis of inversion results. On the longer term, the intention is to avoid that our initial design choices restrict the implementation of future experiments.

For this initial paper, we have performed regional CO_2 inversions, intentionally using a rather 'classical' inversion design to ensure comparability with other similar setups and to have a reference of comparison for future inversions, but also because it enablesd us to focus on the technical robustness of the code. The transport is performed by a script which relies on pre-computed FLEXPART observation footprints, and on background concentrations pre-computed with the global coarse resolution TM5-4DVAR inverse model (although technically, nothing limits the use of alternative models to compute these footprints and background concentrations). This TM5-FLEXPART setup replicates the 2-step TM3-STILT inversion proposed by Rödenbeck et al. (2009).

Although the inversion setup lacks the maturity of established systems, it offers promising computational performances and the results suggest interesting scientific questions regarding the capacity of regional inversion systems to constrain the annual budget of CO_2, and point to specific improvements of the inversion approach, which will be implemented in the near future, e.g. the optimization of fluxes at a higher temporal resolution. On the longer term, the aim is to use LUMIA as a platform for testing innovative inversion approaches (multiple transport models, use of satellite data, multi-tracer inversions, optimisation of vegetation model parameters (CCDAS), etc.). The code corresponding to the inversions in this manuscript is provided for the research community on https://lumia.nateko.lu.se, and the access to the git server can be granted on demand.

Code availability. The LUMIA source code used in this paper as well as updates can be downloaded from the lumia website: https://lumia.nateko.lu.se

Author contributions. G.M. and M.S designed the experiments and G.M. developped the code and performed the simulations. G.M. prepared the manuscript and M.S. provided corrections and suggestions for improvements.

Acknowledgements. G.M. has been funded by the Swedish Research Council project 'Development of regional ecosystem-atmosphere models assimilating the ICOS data for a European-scale intercomparison of net CO2 fluxes - Eurocom' (DNR 349-2014-6576). The research is part of three Swedish strategic research areas: ModElling the Regional and Global Earth system (MERGE), the e-science collaboration (eSSENCE), and Biodiversity and Ecosystems in a Changing Climate (BECC).

We thank Michael Mischurow for providing the LPJ-GUESS net ecosystem exchange data, Philippe Peylin for providing the ORCHIDEE NEE fluxes, and Greet Janssens-Maenhout for providing the fossil fuel product. We also thank the National Supercomputer Centre at Linköping University (part of the Swedish National Infrastructure for Computing, SNIC) for providing computer resources for the project.

We thank the FLEXPART and TM5 developers for providing the transport models source codes.

Finally, we thank all the observation data providers cited in Table 1 for providing the observations used in this paper, and we thank NOAA/GMD for collecting and distributing the data in obspack format.

References

BP: Statistical Review of World Energy 2016, Tech. rep., http://www.bp.com/content/dam/bp/excel/energy-economics/statistical-review-2016/bp-statistical-review-of-world-energy-2016-workbook.xls, 2016.

Basu, S., Guerlet, S., Butz, A., Houweling, S., Hasekamp, O., Aben, I., Krummel, P., Steele, P., Langenfelds, R., Torn, M., Biraud, S., Stephens, B., Andrews, A., and Worthy, D.: Global CO2 fluxes estimated from GOSAT retrievals of total column CO2, Atmospheric Chemistry and Physics, 13, 8695–8717, https://doi.org/10.5194/acp-13-8695-2013, http://www.atmos-chem-phys.net/13/8695/2013/, 2013.

Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T., and Ciais, P.: A European summertime CO2 biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, Journal of Geophysical Research: Atmospheres, 116, n/a–n/a, https://doi.org/10.1029/2011JD016202, http://doi.wiley.com/10.1029/2011JD016202, 2011.

Chevallier, F., Fisher, M., Peylin, P., Serrar, S., Bousquet, P., Bréon, F.-M., Chédin, A., and Ciais, P.: Inferring CO2 sources and sinks from satellite observations: Method and application to TOVS data, Journal of Geophysical Research, 110, D24 309, https://doi.org/10.1029/2005JD006390, 2005.

Ciais, P., Crisp, D., Denier van der Gon, H., Engelen, R., Janssens-Maenhout, G., Heimann, M., Rayner, P., and Scholze, M.: Towards a European Operational Observing System to Monitor Fossil CO2 emissions - Final Report from the expert group, vol. 19, European Commission, Copernicus Climate Change Service, https://www.copernicus.eu/sites/default/files/2018-10/CO2{_}Report{_}22Oct2015.pdf, 2015.

Ciattaglia, L., Cundari, V., and Colombo, T.: Further measurements of atmospheric carbon dioxide at Mt. Cimone, Italy: 1979–1985, Tellus B, 39 B, 13–20, https://doi.org/10.1111/j.1600-0889.1987.tb00266.x, 1987.

Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, Q. J. R. Mereorol Soc, 120, 1367–1387, http://onlinelibrary.wiley.com/doi/10.1002/qj.49712051912/abstract, 1994.

Denier van der Gon, H. D., Hendriks, C., Kuenen, J., Segers, A., and Visschedijk, A.: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns, in: EU FP7 MACC deliverable report D_D-EMIS_1.3, https://atmosphere.copernicus.eu/sites/default/files/2019-07/MACC{_}TNO{_}del{_}1{_}3{_}v2.pdf, 2011.

Dlugokencky, E., Mund, J. M., Crotwell, A., Crotwell, M., and Thoning, K.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2018, Version: 2019-07, https://doi.org/https://doi.org/10.15138/wkgj-f215, 2019.

Errico, R. M.: What Is an Adjoint Model ?, Bulletin of the American Meteorological Society, 78, 2577–2591, 1997.

Farquhar, G. D., von Caemmerer, S., and Berry, J. A.: A biochemical model of photosynthetic CO2 assimilation in leaves of C3 species, Planta, 149, 78–90, https://doi.org/10.1007/BF00386231, https://doi.org/10.1007/BF00386231, 1980.

Ganesan, A. L., Manning, A. J., Grant, A., Young, D., Oram, D. E., Sturges, W. T., Moncrieff, J. B., and O'Doherty, S.: Quantifying methane and nitrous oxide emissions from the UK and Ireland using a national-scale monitoring network, Atmospheric Chemistry and Physics, 15, 6393–6406, https://doi.org/10.5194/acp-15-6393-2015, http://www.atmos-chem-phys.net/15/6393/2015/, 2015.

Giglio, L., Randerson, J. T., and Van Der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), Journal of Geophysical Research: Biogeosciences, 118, 317–328, https://doi.org/10.1002/jgrg.20042, 2013.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y. H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J.,

Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C. W.: Towards robust regional estimates of annual mean {CO}_2 sources and sinks, Nature, 415, 626–630, 2002.

Hammer, S., Glatzel Mattheier, H., Müller, L., Sabasch, M., Schmidt, M., Schmitt, S., Schönherr, C., Vogel, F., Worthy, D., and Levin, I.: A gas chromatographic system for high-precision quasi-continuous atmospheric measurements of CO2, CH4, N2O, SF6, CO and H2., 2008.

Haszpra, L., Barcza, Z., Bakwin, P., Berger, B., Davis, K. J., and Weidinger, T.: Measuring system for the long-term monitoring of biosphere/atmosphere exchange of carbon dioxide, Journal of Geophysical Research Atmospheres, 106, 3057–3069, https://doi.org/10.1029/2000JD900600, 2001.

Hatakka, J., Aalto, T., Aaltonen, V., Aurela, M., Hakola, H., Komppula, M., Laurila, T., Lihavainen, H., Paatero, J., Salminen, K., and Viisanen, Y.: Overview of the atmospheric research activities and results at Pallas GAW station, Boreal Environment Research, 8, 365–383, 2003.

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

Huijnen, V., Williams, J., Van Weele, M., Van Noije, T., Krol, M., Dentener, F., Segers, A., Houweling, S., Peters, W., De Laat, J., Boersma, F., Bergamaschi, P., Van Velthoven, P., Le Sager, P., Eskes, H., Alkemade, F., Scheele, R., Nédélec, P., and Pätz, H. W.: The global chemistry transport model TM5: Description and evaluation of the tropospheric chemistry version 3.0, Geoscientific Model Development, 3, 445–473, https://doi.org/10.5194/gmd-3-445-2010, 2010.

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., Fisk, J. P., Hibbard, K., Houghton, R. A., Janetos, A., Jones, C. D., Kindermann, G., Kinoshita, T., Klein Goldewijk, K., Riahi, K., Shevliakova, E., Smith, S., Stehfest, E., Thomson, A., Thornton, P., van Vuuren, D. P., and Wang, Y. P.: Harmonization of land-use scenarios for the period 1500-2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, Climatic Change, 109, 117–161, https://doi.org/10.1007/s10584-011-0153-2, 2011.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J. G. J., Peters, J. A. H. W., van Aardenne, J. A., Monni, S., Doering, U., Petrescu, A. M. R., Solazzo, E., and Oreggioni, G. D.: EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012, Earth System Science Data, 11, 959–1002, https://doi.org/10.5194/essd-11-959-2019, https://www.earth-syst-sci-data.net/11/959/2019/, 2019.

Jones, A., Thomson, D., Hort, M., and Devenish, B.: The U.K. Met Office's next generation atmospheric dispersion model, NAME III, in: Air pollution modeling and its applications XVII (Proceedings of the 27th NATO/CCMS International Technical Meeting on Air Pollution Modelling and it's Application), edited by Borrego, C. and A.-L, N., pp. 580–589, Springer, 2007.

Kaminski, T., Knorr, W., Schürmann, G., Scholze, M., Rayner, P. J., Zaehle, S., Blessing, S., Dorigo, W., Gayler, V., Giering, R., Gobron, N., Grant, J. P., Heimann, M., Hooker-Stroud, A., Houweling, S., Kato, T., Kattge, J., Kelley, D., Kemp, S., Koffi, E. N., Köstler, C., Mathieu, P. P., Pinty, B., Reick, C. H., Rödenbeck, C., Schnur, R., Scipal, K., Sebald, C., Stacke, T., Van Scheltinga, A. T., Vossbeck, M., Widmann, H., and Ziehn, T.: The BETHY/JSBACH Carbon Cycle Data Assimilation System: Experiences and challenges, Journal of Geophysical Research: Biogeosciences, 118, 1414–1426, https://doi.org/10.1002/jgrg.20118, 2013.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.: Technical Note: Atmospheric CO2 inversions on the mesoscale using data-driven prior uncertainties: methodology and system evaluation, Atmospheric Chemistry and Physics, 18, 3027–3045, https://doi.org/10.5194/acp-18-3027-2018, https://www.atmos-chem-phys.net/18/3027/2018/, 2018.

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 Biogeochemical Cycles, 19, 1–33, https://doi.org/10.1029/2003GB002199, 2005.

Lanczos, C.: An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators, Journal Of Research Of The National Bureau Of Standards, 45, 255–282, 1950.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-i., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster, U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth System Science Data, 10, 2141–2194, https://doi.org/10.5194/essd-10-2141-2018, https://www.earth-syst-sci-data.net/10/2141/2018/, 2018.

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

Meirink, J. F., Bergamaschi, P., and Krol, M. C.: Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion, Atmospheric Chemistry and Physics, 8, 6341–6353, 2008.

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived greenhouse gases, Atmospheric Chemistry and Physics Discussions, pp. 1–27, https://doi.org/10.5194/acp-2016-800, http://www.atmos-chem-phys-discuss.net/acp-2016-800/{#}discussion, 2016.

Michalak, A. M., Randazzo, N. A., and Chevallier, F.: Diagnostic methods for atmospheric inversions of long-lived greenhouse gases, Atmospheric Chemistry and Physics, 17, 7405–7421, https://doi.org/10.5194/acp-17-7405-2017, https://www.atmos-chem-phys.net/17/7405/2017/, 2017.

NOAA Carbon Cycle Group ObsPack Team: Multi-laboratory compilation of atmospheric methane data for the period 1957-2017; obspack_ch4_1_GLOBALVIEWplus_v1.0_2019_01_08;, https://doi.org/https://doi.org/10.25925/20190108, https://www.esrl.noaa.gov/gmd/ccgg/obspack/data.php?id=obspack{_}ch4{_}1{_}GLOBALVIEWplus{_}v1.0{_}2019-01-08, 2019.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Pétron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker., Proceedings of the National Academy of Sciences of the United States of America, 104, 18 925–18 930, https://doi.org/10.1073/pnas.0708986104, http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2141884{&}tool=pmcentrez{&}rendertype=abstract, 2007.

Peylin, P., Law, R. M., Gurney, K. R., Chevallier, F., Jacobson, A. R., Maki, T., Niwa, Y., Patra, P. K., Peters, W., Rayner, P. J., Rödenbeck, C., Van Der Laan-Luijkx, I. T., and Zhang, X.: Global atmospheric carbon budget: Results from an ensemble of atmospheric CO2 inversions, Biogeosciences, 10, 6699–6720, https://doi.org/10.5194/bg-10-6699-2013, 2013.

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N., Cassiani, M., Eckhardt, S., Arnold, D., Morton, D., Thompson, R. L., Groot Zwaaftink, C. D., Evangeliou, N., Sodemann, H., Haimberger, L., Henne, S., Brunner, D., Burkhart, J. F., Fouilloux, A., Brioude, J., Philipp, A., Seib-

ert, P., and Stohl, A.: The Lagrangian particle dispersion model FLEXPART version 10.3, Geoscientific Model Development Discussions, pp. 1–67, https://doi.org/10.5194/gmd-2018-333, https://www.geosci-model-dev-discuss.net/gmd-2018-333/, 2019.

Rayner, P. J., Michalak, A. M., and Chevallier, F.: Fundamentals of Data Assimilation applied to biogeochemistry, Atmospheric Chemistry and Physics Discussions, pp. 1–32, https://doi.org/10.5194/acp-2018-1081, https://www.atmos-chem-phys-discuss.net/acp-2018-1081/, 2018.

Rödenbeck, C., Gerbig, C., Trusilova, K., and Heimann, M.: A two-step scheme for high-resolution regional atmospheric trace gas inversions based on independent models, Atmospheric Chemistry and Physics Discussions, 9, 1727–1756, https://doi.org/10.5194/acpd-9-1727-2009, 2009.

Rödenbeck, C., Keeling, R. F., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean pCO2 and sea-Air CO2 flux variability from an observation-driven ocean mixed-layer scheme, Ocean Science, 9, 193–216, https://doi.org/10.5194/os-9-193-2013, 2013.

Rozanski, K., Necki, J., Chmura, L., Sliwka, I., Zimnoch, M., Bielewski, J., Galkowski, M., Bartyzel, J., and Rosiek, J.: Anthropogenic changes of CO2, CH4, N2O, CFCl3, CF2Cl2, CCl2FCClF2, CHCl3, CH3CCl3, CCl4, SF6 and SF5CF3 mixing ratios in the atmosphere over southern Poland, Geological Quarterly, 58, https://doi.org/10.7306/gq.1163, 2014.

Schaefer, K., Collatz, G. J., Tans, P., Denning, A. S., Baker, I., Berry, J., Prihodko, L., Suits, N., and Philpott, A.: Combined simple biosphere/carnegie-ames-stanford approach terrestrial carbon cycle model, Journal of Geophysical Research: Biogeosciences, 113, 1–13, https://doi.org/10.1029/2007JG000603, 2008.

Schmidt, M.: The Schauinsland CO$_2$ record: 30 years of continental observations and their implications for the variability of the European CO$_2$ budget, Journal of Geophysical Research, 108, 1–7, https://doi.org/10.1029/2002jd003085, 2003.

Seibert, P. and Frank, a.: Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode, Atmospheric Chemistry and Physics, 4, 51–63, https://doi.org/10.5194/acp-4-51-2004, 2004.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S. D., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P. E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J. G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S. L., Le Quéré, C., Smith, B., Zhu, Z., and Myneni, R.: Recent trends and drivers of regional sources and sinks of carbon dioxide, Biogeosciences, 12, 653–679, https://doi.org/10.5194/bg-12-653-2015, https://www.biogeosciences.net/12/653/2015/, 2015.

Smith, B., Wärlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., and Zaehle, S.: Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model, Biogeosciences, 11, 2027–2054, https://doi.org/10.5194/bg-11-2027-2014, 2014.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., and Mukai, H.: The CO2 release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmospheric Chemistry and Physics, 11, 6855–6870, https://doi.org/10.5194/acp-11-6855-2011, https://www.atmos-chem-phys.net/11/6855/2011/, 2011.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, in: IPCC, 2013, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Stohl, A., Sodemann, H., Eckhardt, S., Frank, A., Seibert, P., and Wotawa, G.: The Lagrangian particle dispersion model FLEXPART version 8.2, Tech. rep., 2010.

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., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., 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 pCO2, and net sea–air CO2 flux over the global oceans, Deep Sea Research Part II: Topical Studies in Oceanography, 56, 554–577, https://doi.org/10.1016/J.DSR2.2008.12.009, https://www.sciencedirect.com/science/article/pii/S0967064508004311, 2009.

Thompson, R. L., Stohl, A., Zhou, L. X., Dlugokencky, E., Fukuyama, Y., Tohjima, Y., Kim, S. Y., Lee, H., Nisbet, E. G., Fisher, R. E., Lowry, D., Weiss, R. F., Prinn, R. G., O'Doherty, S., Young, D., and White, J. W.: Methane emissions in East Asia for 2000-2011 estimated using an atmospheric Bayesian inversion, Journal of Geophysical Research, 120, 4352–4369, https://doi.org/10.1002/2014JD022394, 2015.

Turner, a. J. and Jacob, D. J.: Balancing aggregation and smoothing errors in inverse models, Atmospheric Chemistry and Physics, 15, 7039–7048, https://doi.org/10.5194/acp-15-7039-2015, http://www.atmos-chem-phys.net/15/7039/2015/, 2015.

Uglietti, C., Leuenberger, M., and Brunner, D.: European source and sink areas of CO2 retrieved from Lagrangian transport model interpretation of combined O2 and CO2 measurements at the high alpine research station Jungfraujoch, Atmospheric Chemistry and Physics, 11, 8017–8036, https://doi.org/10.5194/acp-11-8017-2011, 2011.

van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: A single gas chromatograph for accurate atmospheric mixing ratio measurements of CO₂, CH₄, N₂O, SF₆ and CO, Atmospheric Measurement Techniques Discussions, 2, 1321–1349, https://doi.org/10.5194/amtd-2-1321-2009, 2009.

Van Der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997-2016, Earth System Science Data, 9, 697–720, https://doi.org/10.5194/essd-9-697-2017, 2017.

Vermeulen, a. T., Hensen, a., Popa, M. E., Van Den Bulk, W. C. M., and Jongejan, P. a. C.: Greenhouse gas observations from Cabauw Tall Tower (1992-2010), Atmospheric Measurement Techniques, 4, 617–644, https://doi.org/10.5194/amt-4-617-2011, 2011.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505–7514, https://doi.org/10.1002/2014WR015638.Received, 2014.

Wilson, P.: Insight into the Carbon Cycle from Continuous Measurements of Oxygen and Carbon Dioxide at Weybourne Atmospheric Observatory , UK, Ph.D. thesis, University of East Anglia, 2012.

Yver, C., Schmidt, M., Bousquet, P., and Ramonet, M.: Measurements of molecular hydrogen and carbon monoxide on the Trainou tall tower, Tellus, Series B: Chemical and Physical Meteorology, 63, 52–63, https://doi.org/10.1111/j.1600-0889.2010.00520.x, 2011.