https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

# Impact of atmospheric transport on $CO_2$ flux estimates derived from the atmospheric tracer inversions

Saqr Munassar[1,2], Guillaume Monteil[3], Marko Scholze[3], Ute Karstens[4], Christian Rödenbeck[1], Frank-Thomas Koch[1,5], Kai U. Totsche[6], and Christoph Gerbig[1]

[1]Department of Biogeochemical Signals, Max-Planck Institute for Biogeochemistry, Jena, Germany
[2]Department of Physics, Faculty of Sciences, Ibb University, Ibb, Yemen
[3]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden
[4]ICOS Carbon Portal at Lund University, Lund, Sweden
[5]Meteorological Observatory Hohenpeissenberg, Deutscher Wetterdienst, Hohenpeißenberg, Germany
[6]Institute of Geoscience, Friedrich Schiller University, Jena, Germany

*Correspondence to*: Saqr Munassar (smunas@bgc-jena.mpg.de)

**Abstract.** We present an analysis of atmospheric transport impact on estimating $CO_2$ fluxes using two atmospheric inversion systems (CarboScope Regional (CSR) and LUMIA) over Europe for 2018. The main focus of this study is to quantify the dominant drivers of spread amid $CO_2$ estimates derived from atmospheric tracer inversions. The Lagrangian transport models STILT and FLEXPART were used to assess the impact of mesoscale transport. The impact of lateral boundary conditions for $CO_2$ was assessed by applying the global transport models TM3 and TM5. $CO_2$ estimates calculated with an ensemble of eight inversions differing in the regional and global transport models, as well as the inversion systems show a relatively large spread for the annual domain wide flux ranging between -0.72 and 0.20 PgC $yr^{-1}$ with a mean estimate of -0.29 PgC. The largest discrepancies resulted from varying the mesoscale transport model, which amounted to a difference of 0.51 (PgC $yr^{-1}$), in comparison with 0.23 and 0.10 (PgC $yr^{-1}$) that resulted from the far-field contributions and the inversion systems, respectively. Additionally, varying the mesoscale transport caused large discrepancies in spatial and temporal patterns, while changing the lateral boundary conditions lead to more homogeneous spatial and temporal impact. We further investigated the origin of the discrepancies between transport models. The meteorological forcing parameters (forecasts versus reanalysis obtained from ECMWF data products) used to drive the transport models are responsible for a small part of the differences in $CO_2$ estimates, but the largest impact seems to come from the models themselves. Although a good convergence in the differences between the inversion systems was achieved by applying a strict protocol of using identical priors, and atmospheric datasets, there was a non-negligible impact arising from applying a different inversion system. Specifically, the choice of prior error structure accounted for a large part of system-to-system differences.

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

## 1 Introduction

Inverse modeling has been increasingly used to infer surface-atmosphere fluxes of carbon dioxide ($CO_2$), from observations of dry mole fractions made at spatiotemporal points across an observational network (Enting and Newsam, 1990; Bousquet et al., 1999). An atmospheric transport model is used to relate the measured tracer concentration to its possible sources and sinks, which are adjusted in order to fit the modeled to observed data. However, inaccuracies in the representation of the real atmospheric dynamics by the transport model lead to uncertainties in $CO_2$ flux estimates. Transport model errors can emerge from both simplified parameterizations of real physics and model parameters themselves (Engelen, 2002).

An atmospheric transport model relies on a mesoscale representation of air mass movements, which cannot completely reproduce the observed fine scale variability of tracer concentration, leading to the so-called representation error. Additionally, inversions cannot solve for fluxes at a lower spatial and temporal resolutions than that of their transport model, which leads to aggregation errors (Kaminski et al., 2001). Likewise, imperfection in representing mixed layer heights results in errors in the modeled tracer concentration. Gerbig et al. (2008) found large discrepancies in derived mixing heights between meteorological analysis from ECMWF and radiosonde data, which reached about 40% for the daytime and about 100% for the nocturnal boundary layer, correspondingly impacting simulated $CO_2$. Tolk et al. (2008) also found meteorology to be a key driver of representation error, which varies spatially and temporally. They indicated that a large contribution to representation error is caused by unresolved model topography at coarse spatial resolution during night, but that convective structures, mesoscale circulations, and the variability of $CO_2$ fluxes dominate during day-time.

The accuracy of atmospheric transport models typically depends upon the accuracy of meteorological data, since transport models are forced by meteorological parameters. Deng et al. (2017) found that assimilating meteorological observations such as wind speed and wind direction in transport models significantly improved the model performances achieving an uncertainty reduction of about 50% in wind speed and direction, especially when measurements in the Planetary Boundary Layer (PBL) were assimilated. However, they concluded that the differences in $CO_2$ emissions reached up to 15% at local scale corrections after inversion and were limited to 5% for the total emissions integrated across the regional domain of interest. A further study suggests that uncertainties in both transport and $CO_2$ fluxes contribute equally to the uncertainties in $CO_2$ dry mole fraction simulations, displaying similar temporal and spatial patterns (Chen et al., 2019). Moreover, systematic differences between transport models occur due to discrepancies in representing vertical mixing, horizontal winds, as well as horizontal and vertical resolution of the models (Peylin et al., 2002). Although $CO_2$ fluxes constrained by atmospheric data in the Bayesian inversion framework inherit a dominant spatial and temporal pattern from the atmospheric signal, the a-posteriori still suffers from large spread when using different global and mesoscale transport models (Rivier et al., 2010).

In this study we further investigated the impact of atmospheric transport on estimating biogenic $CO_2$ fluxes using two atmospheric tracer inversion systems at fine spatio-temporal resolutions over the regional domain of Europe (the same domain as used for the EUROCOM project study implemented by Monteil et al. (2020). We present results of posterior NEE estimated from the two inversion systems CarboScope-Regional, CSR, (Kountouris et al., 2018a; Munassar et al., 2022) and

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

LUMIA (Monteil and Scholze, 2021). Both inversions employ pre-computed sensitivities of atmospheric mole fractions to surface fluxes, so-called source-weight functions or "footprints", via two Lagrangian transport models at regional scales, and make use of the two-step inversion approach established by Roedenbeck et al. (2009) to provide the lateral boundary conditions. The regional atmospheric transport models were used at a horizontal resolution of 0.25-degree to assess the impact of mesoscale transport on $CO_2$ estimates. In addition, the resulting impact from using different boundary conditions was assessed. The impact of both global and regional models on the flux estimates was compared through analysing the differences in space and time.

Detailed descriptions about the inversion set-ups, the transport models, and the prior fluxes used in this study are presented in the methods (Section 2). The observational stations that provide $CO_2$ dry mole fraction are described within the methods as well. We introduce the results obtained from eight inversions varying in regional and global transport models using the regional inversions CSR and LUMIA in Section 3. The results are discussed and interpreted through a spatial and temporal analysis of the differences between the elements of inversions in Section 4. Finally, Section 5 presents remarks and conclusions about the impact of the transport models on $CO_2$ estimates in the inverse modeling carried out throughout this study.

## 2 Methods

An atmospheric tracer inversion framework is mainly made up of transport model, data source for boundary conditions (in case of regional inversions), atmospheric dataset of mole fractions, and surface flux fields. In this study, several inversion runs differing in atmospheric transport models are conducted using two tracer inversion systems, CSR and LUMIA (see Table 2). The default CSR inversion system utilizes pre-calculated footprints from the Stochastic Time-Inverted Lagrangian Transport model STILT (Lin et al., 2003) at the regional domain, and the TM3 model at the global scale, applying the two-step scheme inversion approach (Rödenbeck et al., 2009), to provide the far-field contributions to the regional domain. In the default setup of the second inversion system LUMIA, the footprints are pre-calculated using the Lagrangian particle dispersion model FLEXPART (Pisso et al., 2019), and the far-field contributions are calculated using the global transport model TM5 in a separate global inversion run, applying the two-step scheme inversion as well. These default configurations in both the systems constitute the base cases. We strive to restrict the differences in the inversion runs to the atmospheric transport models to outline the impact of atmospheric transport. That is, input data such as measurements of $CO_2$ dry mole fraction and the a-priori fluxes, used as constraints based on Bayes inference, are identical for all runs. We exchangeably make use of the four combinations of transport model components, the regional and global models, in the two inversion systems to evaluate the effect of transport using forward model simulations and inversion runs to quantify the magnitude of differences in the flux space. The inversion setups and implementation are explained in the protocol of comparison (Section 2.6).

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

### 2.1 Inversion Framework

In the following description we remind the reader about the basic principles of the inversion schemes. For detailed information about the mathematical schemes, the reader is referred to (Rödenbeck, 2005) for CSR and to Monteil and Scholze (2021) for LUMIA. Both systems rely on the Bayesian inference that accounts for observations and prior knowledge to regularize the solution of the ill-posed inverse problem where a unique solution does not exist due to the spatial scarcity of observations. Therefore, the optimal state vector (x) is searched for in the Bayesian formalism by minimizing the cost function $\boldsymbol{J(x)}$ that is typically composed of the observational constraint term $\boldsymbol{J_c(x)}$ and the prior flux constraint term $\boldsymbol{J_b(x)}$

$$\boldsymbol{J(x) = J_c(x) + J_b(x)} \quad (1)$$

where

$$\boldsymbol{J_b(x) = \frac{1}{2}\,(x - x_b)^T\,\mathbf{B^{-1}}\,(x - x_b)} \quad (2)$$

$$\boldsymbol{J_c(x) = \frac{1}{2}\,(H(x) - y)^T\,\mathbf{Q^{-1}}\,(H(x) - y)} \quad (3)$$

The prior flux uncertainty defined in the covariance matrix **B** limits the departure of the control vector ($\boldsymbol{x}$) to the prior flux vector ($\boldsymbol{x_b}$). Similarly, the observational constraint is weighted by the observational covariance matrix **Q** that contains the so-called model-data mismatch error, including uncertainty of measurement, representativeness, and transport. This uncertainty is assigned to the diagonal of the matrix **Q** for the respective sites based on the ability of the transport model to represent the atmospheric circulation at such locations. $\boldsymbol{H(x)}$ represents the atmospheric transport operator that determines the relation between fluxes and the modeled tracer concentration, which corresponds spatially and temporally to a given vector of measurements y. Following the gradient descent method, a variational algorithm is applied iteratively to reach the best convergence (global minimum) of the cost function that satisfies the optimal solution of the control vector. The default configurations for constructing the covariance matrices of prior uncertainty are slightly different in CSR and LUMIA. Prior flux uncertainty is assumed to be around 0.47 PgC $yr^{-1}$ over the full domain of Europe derived from the global uncertainty (2.80 PgC) assumed in the CarboScope global inversion for the annual biogenic fluxes. In CSR, this uncertainty is uniformly distributed spatially and temporally in a way that the annual uncertainty aggregated over the entire domain should arrive at the same value. The uncertainty structure assumes a hyperbolic decay function in space (Eq. (4)) and an exponential function (Eq. (5)) for the temporal decay as explained in Kountouris et al. (2015).

$$r(s) = \frac{1}{1 + \frac{s}{ds}} \quad (4)$$

$$r(t) = e^{\frac{-t}{dt}} \quad (5)$$

The correlation length scales ds and dt applied to flux uncertainties are chosen to be 66.4 km spatially and 30 days temporally, respectively, following Kountouris et al. (2018b). The spatial length in the zonal direction is set to be longer than that in the meridional direction by a factor of 2 (anisotropic), owing to larger spatial climate variability in meridional as compared to zonal direction.

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

The spatio-temporal shape of the prior uncertainty in LUMIA is computed in a way that each control vector comprises weekly uncertainty calculated as the standard deviation of NEE based on weekly flux variance; however, LUMIA agrees on the overall annually aggregated flux uncertainty over the entire domain with CSR. A Gaussian function of the spatial correlation decay (Eq. (6)) is applied to the prior uncertainty structure with a length scale of 500 km

$$r(s) = e^{-(\frac{s}{ds})^2} \quad (6)$$

whereas the effective temporal decay was set to 30 days (same as in CSR). Given the difference in the spatial correlation decay of the prior uncertainty, LUMIA is set to draw larger flux corrections in a broader radial area where stations exist following the gaussian decay with a longer length scale compared to the hyperbolic decay in CSR. In turn, the hyperbolic function has a larger impact in the further radial distances than does the Gaussian function, regardless of the longer spatial scale assumed with the Gaussian decay in a factor of around 7.5 in comparison with the hyperbolic decaying function.

### 2.2 Atmospheric transport models

Surface sensitivities are calculated using the STILT (Lin et al., 2003) and FLEXPART (Pisso et al., 2019) models at a horizontal resolution of 0.25-degree and hourly temporal resolution. Both models simulate the transport of air mass via releasing an ensemble of virtual particles at the locations of stations. The virtual particles are transported backward in time and driven by meteorological fields obtained from the European Center for Medium-Range Weather Forecasts (ECMWF). STILT particles are transported 10 days backward in time and forced by forecasting data obtained from the high-resolution implementation of the Integrated Forecasting System (IFS HRES). For the FLEXPART model, in standard operation particles are followed for 15 days driven by ERA-5 reanalysis data. To keep the consistency with STILT footprints, the backward time of FLEXPART footprints was limited to 10 days in the inversions. After such a maximum backward time integration, the particles are assumed to leave the domain, even though a large number of particles are expected to escape after a few days. To better represent air sampling in the mixed layer, day-time observations are considered, except for mountain stations where night-time observations are used instead (Geels et al., 2007). To ensure best mixing conditions, temporal windows were considered for simulating $CO_2$ dry model fractions over stations as explained in Section 2.4 (Table 1). In addition, release heights of particles are taken as the highest sampling level above ground at each measurement site. For high altitude receptors such as mountains, a correction height is used in STILT in a way that the actual elevation of the station can be represented in the corresponding vertical model level (Munassar et al., 2022). In FLEXPART the elevation above sea level is taken as the model sampling height.

### 2.3 A priori and prescribed fluxes

Three components of prior and prescribed surface-to-atmosphere fluxes of $CO_2$ are used here and kept identical in both systems: 1) biogenic terrestrial fluxes, 2) ocean fluxes, and 3) anthropogenic emissions. Prior net terrestrial $CO_2$ exchange fluxes, Net Ecosystem Exchange (NEE), are calculated using the diagnostic biogenic model Vegetation Photosynthesis and

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

Respiration Model (VPRM) (Mahadevan et al., 2008). VPRM calculates NEE over the domain of Europe at hourly temporal and 0.25-degree spatial resolution, and provides a partitioning of the net flux into gross ecosystem exchange (GEE) and ecosystem respiration. Data obtained from remote sensing provided through the MODIS instrument and meteorological parameters from ECMWF derive both quantities of the light-dependent GEE and the light-independent ecosystem respiration. The model parameters were also optimized against eddy covariance data selected within the global FLUXNET site network across Europe in 2007 (Kountouris et al., 2015). For more details on the VPRM model the reader is referred to Mahadevan et al. (2008).

Ocean fluxes are taken from Fletcher et al. (2007) which provide climatological fluxes at 5x4 degrees of spatial resolution, remapped to 0.25-degree to be compatible with the biosphere model fluxes. In addition, anthropogenic emissions are taken from the EDGAR_v4.3 inventory, updated to recent years according to British Petroleum (BP) statistics of fossil fuel consumption, and distributed spatially and temporally based on fuel type, category, and country specific emissions, using the COFFEE approach (Steinbach et al., 2011). The emissions are remapped onto a 0.25-degree spatial grid and to hourly temporal resolution.

Biogenic terrestrial fluxes are optimized in the inversions, while the ocean fluxes and anthropogenic emissions are prescribed, given the better knowledge about their spatial and temporal distribution in comparison with the heterogeneity, variability, and uncertainty of the biogenic fluxes. Moreover, in the absence of observational constraints to help discriminate the contributions from the three categories, we chose to prescribe the ocean fluxes and anthropogenic $CO_2$ emissions. This is also justified by the fact that the observation sites are located in areas where the biospheric flux influence is expected to dominate the $CO_2$ concentration variability, but it means that errors in the fossil or ocean fluxes may have to be compensated by the inversions, and resulting in changes in the posterior NEE.

### 2.4 Observations

Measurements of $CO_2$ dry model fractions are collected through ICOS, NOAA, and pre-ICOS stations across the domain of Europe provided by Drought 2018 Team and ICOS Atmosphere Thematic Centre (doi:10.18160/ERE9-9D85, 2020). In total, datasets from 44 stations are used covering 2018, in which a maximum number of stations is present compared to the other years. Regarding model-mismatch errors, in LUMIA a weekly value of 1.5 ppm is assumed to all sites except for the Heidelberg site where 4 ppm was assumed due to the anthropogenic influence from the neighbourhood. Table 1 denotes the weekly values of uncertainty used in CSR for the corresponding sites. The uncertainty for the surface sites is inflated to 2.5 ppm as a slight difference to LUMIA. The inflation of uncertainty from weekly to hourly values is basically calculated by multiplying weekly errors by $\sqrt{7 \times n}$ (n refers to the number of hours in the daily measurements used in the inversion). The observations are mostly assimilated as hourly continuous measurements, and are taken from the highest level, avoiding large vertical gradients near the surface that are hard to represent in the transport models. Model error in representing observations in the PBL is expected to be largest when the PBL is shallow. Therefore, for most sites, we considered data only when the PBL was expected to be well developed, i.e., during the afternoon, local time (LT). The exception is at high altitude sites,

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

which tend to sample the free troposphere during night (Kountouris et al., 2018a). The assimilated windows are reported in Table 1.

### 2.5 Boundary conditions

Far-field contributions of $CO_2$ concentrations (originating from sources outside of the regional domain) are taken from global inversions. As default set ups of the global runs, the Eulerian transport model TM3 is used in the CarboScope global inversion at 5° (lon) x 4° (lat), while TM5 is used in LUMIA at 6° (lon) x 4° (lat) (Monteil and Scholze, 2021). Both inversion systems apply the two-step scheme inversion, explained in Roedenbeck et al. (2009), in which a global inversion is first used to estimate $CO_2$ fluxes globally (based on observations inside and outside Europe). In a second step, the global transport model is used to estimate the influence of European $CO_2$ fluxes on European $CO_2$ observations. That regional influence is then subtracted from the total concentration, to obtain a time-series of the far-field influence directly at the locations of the observation sites. This prevents introducing biases by passing concentration fields from one model to another. For detailed information about the approach methodology, the reader is referred to Roedenbeck et al. (2009).

### 2.6 Comparison protocol

The results of the study are based on eight variants of inversions differing in global and regional transport models, as well as in inversion systems as explained in Table 2. This implies, the two inversion systems (CSR and LUMIA) make use of two regional transport models (STILT and FLEXPART) and two global transport models (TM3 and TM5), which represent the boundary conditions (background) in the regional runs. Hereafter, the identifier codes (see corresponding column in Table 2) will be used to refer to the individual runs within the inversion ensemble. For instance, to highlight the impact of regional transport models, we compare the inversions that only differ in regional transport models, regardless of the inversion system or boundary conditions used, such as CS3 and CF3 or LS5 and LF5. Similarly, we use the same specifications of transport models (indicated through the identifier codes) for the forward runs to outline the differences in prior concentrations simulated with different transport models. In this case using a different system should not result in discrepancies as long as prior fluxes remain identical. In terms of system-to-system comparison, the impact of flux uncertainty should be taken into account as the error structure of the prior is specific for each inversion system. However, this has been investigated by conducting additional tests in CSR and LUMIA using identical uncertainties with flat shape and Gaussian correlation decay.

## 3 Results

### 3.1 Main inversion results

We present estimates of biosphere-atmosphere fluxes for 2018 over the domain of Europe calculated from an ensemble of eight inversions listed in Table 2. All the inversions agree that the biosphere flux model VPRM (prior) overestimates $CO_2$ uptake during the growing season (Fig. 1, left). Ensemble-averaged annual estimates of posterior fluxes yield -0.29 PgC

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

versus an a-priori of -1.49 PgC (Fig. 1, right). However, the spread among posterior estimates is relatively large and ranges between -0.72 and 0.20 PgC $yr^{-1}$ for the annual estimates. The mean of standard deviations calculated for the monthly estimates over the ensemble of inversions results in ±0.72 (PgC $yr^{-1}$). Larger deviations in the estimates are observed over inversions that vary in regional transport such as CS3 and CF3 (or LS5 and LF5) in comparison with those that are calculated with different boundary conditions and systems. Moreover, STILT-based inversions indicate a larger seasonal amplitude compared to FLEXPART-based inversions. CSR and LUMIA agree quite well in the seasonal amplitude as depicted on the solid and dashed lines (Fig. 1, left).

The spatial distribution of NEE is illustrated in Fig. 2 (first row) showing the spatial distribution of posterior fluxes estimated with the inversions CS3 and LF5 (default set-ups in CSR and LUMIA, respectively), as well as the a-priori NEE. Both CS3 and LF5 exhibit good agreement in predicting a smaller uptake of $CO_2$ compared to the a-priori fluxes. The magnitudes of flux corrections made by the corresponding inversions are depicted in the innovations of fluxes (Fig. 2, second row). Major corrections are mostly made around the observational sites where larger surface influences exist and the prior uncertainty is highly correlated. Posterior NEE suggests a weaker annual uptake of $CO_2$ but also a net release over some areas in central and western Europe.

The innovations in CS3 indicate predominant positive flux corrections throughout the full domain, while LF5 innovations suggest larger positive flux corrections in southern Europe and relatively small (but negative) corrections in northern France, Netherlands, and south-eastern UK. These discrepancies are attributable to the combination of differences in regional transport models, lateral boundaries, and system configurations. In the following sections we will attempt to separate the contributions of differences caused by each driver - i.e., mesoscale transport, lateral boundaries, and system configurations.

### 3.2 Differences in optimized fluxes and forward simulations

In this section, we look at the impact of changing individual components of the inversion system on the inversion results, and we put them in relation to differences in prior model-data mismatches. Differences in the optimized monthly fluxes caused by the different regional transport models, different inversion systems, and different boundary conditions are shown in Fig. 3 (top). Inversions differing in the regional transport models STILT and FLEXPART exhibit the largest differences in posterior fluxes (transport). The discrepancies caused by transport indicate an obvious seasonal pattern, while boundary conditions (background) and inversion systems (system) appear to result in smaller differences that are systematic across all months for the background and more random for the differences related to inversion systems. The magnitude of flux differences is found to be larger in the background compared to inversion systems.

Figure 3, bottom, illustrates the magnitude of differences in forward simulations of $CO_2$ concentrations at hourly time-steps, computed using prior fluxes, averaged over the observational sites distributed across Europe to highlight the impact of transport models on representing the concentrations of $CO_2$ using STILT and FLEXPART models, as well as using different lateral boundary conditions provided from TM3 and TM5. In this case, the forward model runs with STILT and FLEXPART using TM3 boundary conditions (i.e., CS3 and CF3 set-ups) demonstrate quite large differences, while the different

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

boundary conditions show relatively small and consistent differences over the months. Note that these differences preserve the same temporal pattern as the flux differences but with opposite signs. Further diagnostics of model-data mismatches are provided in the supplementary materials indicating the performances of STILT and FLEXPART with respect to the observations using prior and posterior fluxes across the site network at hourly, weekly and yearly time steps (see Fig. 1S and Table 1S).

#### 3.2.1 Impact of mesoscale transport

In this section, we use the inversion results obtained from CS3 and CF3 to analyse the differences in flux estimates resulting from varying regional transport models. The regional transport leads to large spatial discrepancies in annual flux estimates (Fig. 4, "diff: transport"). Using STILT generally leads to predicting a larger release of $CO_2$ in the regional inversions, in particular over central Europe and the UK compared to using FLEXPART. In turn, inversions using FLEXPART suggest less uptake over northern Italy, Switzerland, and south-eastern France. However, this impact refers to a spatial pattern of transport differences that might be caused either by meteorological data or by problematic sites that are hard to represent by transport models. Unlike other areas, this impact is persistent over months, as shown by Fig. 4 ("sd: transport") which shows the standard deviation of monthly differences between CS3 and CF3. In the context of temporal variations, the inversions performed with different regional transport models indicate larger monthly flux variations in comparison with those differing in global models and inversion systems (see Fig. 4, "sd: background" and "sd: system").

The differences between STILT- and FLEXPART-based inversion results for the monthly domain-aggregated estimates (Fig. 3, top) peaked in November and June yielding 2.11 and -1.82 (PgC $yr^{-1}$), respectively. The best agreement between both inversions is realized during the transitional months (August and April) pointing to differences of -0.10 and -0.18 (PgC $yr^{-1}$), respectively. This can be attributed to the decline of the net flux magnitude during these months. Regarding differences in the forward simulations (Fig. 3, bottom), the absolute difference ranges from 0.39 to 4.37 (ppm) computed for the monthly means throughout all the sites. The largest differences are obtained during November and May with -4.37 and 3.60 (ppm), respectively. On the other hand, the smallest differences were found to be -0.39, -0.42, and 0.56 (ppm) during September, April, and August, respectively. These findings suggest a maximum impact of the mesoscale transport during the growing season and winter, while the impact converges to its minima during transitional months such as May and September. These differences in prior concentration are temporally consistent with those seen in the posterior fluxes.

Figure 5 shows the spatial flux differences together with differences in prior concentrations simulated using STILT and FLEXPART during June and December. The prior concentrations were calculated at hourly time steps at the observational sites based on prior fluxes. Noteworthy, the differences in NEE, to a large extent, agree in their spatial patterns with the differences in prior concentrations calculated over the station network. Positive differences in simulated prior concentrations correspond to negative differences in optimized fluxes due to the fact that greater surface sensitivities will result in small flux corrections in the flux space, and vice versa. In this case, STILT computes higher surface sensitivities than FLEXPART in June, therefore the CS3 inversion needs to adjust less the prior fluxes to compensate for this. On the contrary, a weaker

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

uptake is suggested by STILT inversion during December over Europe compared to the FLEXPART inversion, except for some areas around northern Italy and south-eastern France. Meanwhile, there are notably particular areas such as western Europe during June and northern Europe during December in which the differences in posterior fluxes and prior concentrations exhibit opposite signs.

#### 3.2.2 Impact of lateral boundary condition

To highlight the impact arising from using different far field contributions on the estimates of regional NEE, Fig. 4 ("diff: background") shows the differences between inversions that differ in the global transport models (TM3 versus TM5). The contributions of lateral concentration appear to have a homogeneous impact across the regional domain. In comparison with mesoscale transport, a smaller impact on total flux estimates arises from varying lateral boundary conditions (Fig. 1, CS3 vs CS5; LS3 vs LS5). Using TM3 suggests greater influence of lateral $CO_2$ concentrations than TM5 does. As a result, the regional flux signal is adjusted to larger mixing ratios when using TM5 over TM3. The standard deviation of the monthly posterior fluxes obtained from CS3-CS5 (Fig. 4,"sd: background") denotes flat temporal variations throughout all the grid-cells.

From the perspective of temporal differences, using different lateral boundaries shows a smaller impact on optimized fluxes aggregated over the entire domain of Europe compared to the mesoscale transport impact as illustrated in the difference between the two inversions CS3 and CS5 (Fig. 3). It can be seen from this impact that larger flux corrections are suggested with the TM5-based inversion compared to TM3-based inversion. Discrepancies in the monthly posterior fluxes between CS3 and CS5 inversions amount to a range of 0.11 and 0.64 (PgC $yr^{-1}$) absolute differences with a mean of 0.40 (PgC $yr^{-1}$). These differences are obviously smaller than the differences in the mesoscale transport inversions but notably larger than the differences arising when replacing the inversion system.

Unlike the regional transport model impact, the boundary condition impact does not lead to seasonality in the differences of $CO_2$ concentrations (Fig 3, bottom), in alignment with the differences in fluxes (Fig. 3, top). The differences are also much smaller than those resulting from swapping the regional transport models, with monthly mean differences ranging from 0.17 to 0.93 (ppm), with a mean of 0.55 (ppm). That way the posterior contribution of regional fluxes to $CO_2$ concentrations increases to compensate for the lower far-field influence. This is mostly the case for stations located inside the regional domain apart from boundaries.

#### 3.2.3 Impact of inversion systems

The inversion systems show the smallest impact on flux estimates at annual and monthly scales (Fig. 1). The absolute monthly differences range between 0.06 and 0.56 (PgC $yr^{-1}$) with a mean of 0.15 (PgC $yr^{-1}$) as shown in Fig. 3 ("system"). This demonstrates the smallest differences amid inversions in comparison with the transport and lateral boundary differences that yielded absolute means of 1.27 and 0.40 (PgC $yr^{-1}$), respectively. The differences peaked during May, June and November, while during the rest of the year the differences remained rather small. The spatial differences shown in Fig. 4

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

("diff: system") alter between positive and negative differences over the domain (but these tend to compensate when aggregating the flux estimates over the full domain). It should be noted that the inversion systems mainly differ in the definition of the shape and structure of the prior uncertainty. Therefore, applying different structure and magnitude of prior flux uncertainty in the inversions may inflate the error in $CO_2$ flux estimates over the underlying regions in the domain, in particular if the spatial differences do not cancel out. In addition, the corresponding standard deviations of monthly estimates ("sd: system") show large temporal variations, in particular over areas that have large spatial differences. The spatial results indicate a non-negligible impact of inversion systems, especially at national and subnational scales. The specification of the control vector following the construction of covariance matrices that devises the flux correction is responsible for part of the system differences.

## 4 Discussion

The inversions generally predict weaker annual uptake of NEE over the entire domain of Europe than does the prior biosphere flux model. Seasonality of posterior NEE is largely derived from the atmospheric signal; however, the seasonal amplitude demonstrates large differences due mainly to the transport. Given the identical priors and observational datasets used in the inversions, mesoscale transport leads to large discrepancies in posterior fluxes owing to differences in regional transport models. Hence, any error in the atmospheric transport is translated into posterior fluxes as flux corrections. For instance, CSR inversions suggest an annual $CO_2$ flux budget of -0.20 PgC with STILT and -0.72 PgC with FLEXPART, indicating a difference of 0.51 PgC in the annual flux budget. A better illustration of NEE differences can be obtained when comparing June and December estimates among the different inversions, so as to avoid the compensating effect of $CO_2$ seasonality. The differences arising from the mesoscale transport during June and December were found to be -1.82 and 1.75 PgC $yr^{-1}$, respectively (see Fig. 3). This finding is in agreement with Schuh et al. (2019) indicating the large impact of mesoscale transport on estimating $CO_2$ fluxes. The transport also showed a large impact on flux seasonality leading to a difference of 49% relative to the mean seasonal cycle. However, Schuh et al. (2019) found smaller differences amounting to about 10-15% of the mean seasonal cycle. In a relevant study, Deng et al. (2017) estimated $CO_2$ emissions using four different setups of meteorological data assimilations in the transport model and found a difference of about 15% in the local corrections although the error reduction of wind direction and wind speed was about 50% due to data assimilation, while the difference over the entire domain was confined to less than 5%. These results finding refer to the limited impact of meteorological data.

The large number of stations within central and western Europe lead to a strong observational constraint that is reflected in the spatial optimized fluxes over that area. Therefore, large spatial differences between the inversions are pronounced around areas where stations exist, precisely for grid cells that have non-zero footprints. A stronger flux gradient in the meridional

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

direction is noticed in posterior flux maps of CSR compared to the zonal direction due to the anisotropic spatial correlation length of prior error structure. The large temporal variations indicate a systematic error that possibly arises from the transport models themselves as well as from meteorological forcing data, as will be discussed.

Forward runs conducted with STILT and FLEXPART to sample $CO_2$ concentrations over the station network point out the spatial difference between the models in simulating the concentration at the locations of sites (Fig. 5). Since all runs utilize identical prior fluxes, it does not matter for the differences whether the prior fluxes were precise enough to reproduce the true concentration or not. The monthly average differences over all stations in 2018 reached up to 4.37 (ppm). Geels et al. (2007) found even a larger spread up to 10 (ppm) calculated with five transport models over ten stations distributed across Europe. The notably large difference reported in that study is likely attributed to the large discrepancies in the model configurations, especially regarding the horizontal resolution and vertical levels used. The prescribed configurations applied in STILT and FLEXPART lead to a reasonably consistent representation of the atmospheric variability at synoptic and diurnal timescales. One important difference between STILT and FLEXPART is that the STILT model has larger sensitivities during summer than FLEXPART, while the opposite holds true during winter. However, exceptions are realized around some individual sites such as Weybourne (WAO) in the UK and Ispra (IPR) in Italy, indicating either difficult terrains that cannot be well represented by models at such sites or real synoptic features that are resolved by one model but not the another. The differences in forward runs are inversely manifested in the posterior flux differences as large surface sensitivities result in less flux corrections to be adjusted by the inversion at the expense of the true fluxes, and vice versa. The differences appeared to be larger during the months of growing season and winter following the seasonal amplitude of $CO_2$ in comparison with the rest of months.

**Drivers of STILT-FLEXPART differences**

Although STILT and FLEXPART are run at the same spatio-temporal resolution employing similar schemes to parametrize the atmospheric motion unresolved by meteorological forcing data such as turbulence, and similar diagnostics to determine mixing heights, they still exhibit large spatial and temporal differences. Due to the fact that deep mixed layer heights dilute the tracer concentration owing to the radiative heating process taking place during day-time, the surface sensitivity and atmospheric stratification are expected to decrease accordingly. Thus, a logical assumption is that large values of STILT footprints during the growing season, as supported by the simulations of forward model runs, would probably be caused by a shallower mixing height in comparison with FLEXPART footprints, while FLEXPART footprints demonstrate larger values during winter for the same reason. However, the relationship between the differences of prior concentrations and the differences of mixed layer heights computed with the two models do not show a striking correlation, as can be seen from the scatter plots of the differences in Fig. 6. This finding concludes that the discrepancies in representing mixed layer heights do not explain the major differences in the $CO_2$ simulations nor the differences in footprints.

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

The second assumption was that differences in the forcing data of meteorology products might lead to the discrepancies in both models, given that STILT uses meteorological parameters from IFS HRES, while FLEXPART uses ERA-5 reanalysis. Results in Fig. 7, ”meteo”, indicate that using different meteorological data results in pronounced differences when the FLEXPART model was forced by operational forecast data instead of ERA-5 reanalysis. These differences notably occur during the time of net $CO_2$ release corresponding to quite small differences during the time of growing season. This, however, only explains a small part of the overall differences (shown in Fig. 7, “base”) that dominate all the months except August and September. In a previous study, Liu et al. (2011) concluded that uncertainties in meteorological fields lead to a significant contribution to the total transport error, as well as to an underestimation of the vertical turbulent mixing even when the same circulation model and mixing parameterizations were used to reconstruct vertical mixing from a single meteorological analysis. In our analysis, one should take into consideration that both operational forecasts and ERA-5 are ECMWF products and thus are expected not to significantly diverge.

Furthermore, we tested the possible impact of surface layer heights (the height up to which particles are sensitive to the fluxes) that may affect the particle dispersion, provided that STILT relies on the assumption of defining the surface layer as a half of the mixed layer height, while in FLEXPART it is defined as a fixed height of 100 m (these are default configurations of the models). In this experiment, STILT was run with a surface layer height of 100 m, so that the impact of the surface layer on $CO_2$ simulations is outlined by the comparison with a run using the default configurations of STILT. When running STILT with a surface layer of 100 m, the differences with the default STILT run (Fig. 7, “s_layer”) were found to be quite small and, therefore, can be negligible in both magnitude and temporal pattern compared to the overall differences. However, varying the models (STILT and FLEXPART are driven by IFS data and using a surface layer of 100 m) lead to the largest differences, in particular during the growing season months and winter months as seen from Fig. 7 (“model”). As a result, model parameterizations are expected to be the main driver of differences. It is clearly noticeable that the overall differences (default configurations of the models) combine the underlying differences of “model”, “meteo”, and “s_layer”, and are yielded as the arithmetic summation of this partitioning. Therefore, the differences in $CO_2$ simulations should considerably be dealt with as a model-model error. A further study would be needed to investigate the main drivers of these discrepancies in both STILT and FLEXPART for further model developments.

**Impact of lateral boundaries and systems**

In terms of far field contribution impact, smaller differences of $CO_2$ estimates are observed when varying the global transport models compared to the mesoscale transport. The differences in posterior fluxes arising from boundary conditions amount to 0.40 (PgC $yr^{-1}$) during June and 0.11 (PgC $yr^{-1}$) during December. This impact was found to be consistent in space and time, with coherent deviation over months, and is not therefore expected to affect the seasonal and interannual variability. Unlike the regional transport model error, this kind of error can be thought of as a bias in dry mole fractions, which may be dealt

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

with in the inversions as a constant correction, potentially site-specific correction. The differences between far field contributions suggest that using TM3 leads to smaller regional signal of mole fractions used to constrain fluxes in comparison with TM5. The consistency of the lateral boundary impact over time and space is in agreement with results of lateral boundary uncertainties assessed in Chen et al. (2019) using four different global transport models, albeit over a different domain.

Using a different inversion system results in non-negligible differences in $CO_2$ flux estimates, albeit smaller than that resulting from transport and lateral boundaries. This impact is however dependent upon system configurations, in particular the way how the flux uncertainty is prescribed. For example, LS3 predicts -6.416 and 2.387 (PgC $yr^{-1}$) during June and December, respectively, which is higher than CS3 estimates by 0.325 and 0.067 (PgC $yr^{-1}$). Generally, LUMIA predicts slightly larger $CO_2$ releases compared to CSR, which is partially attributed to differences in uncertainties assumed in both systems. The impact of uncertainty definition is quantitatively assessed through utilizing identical uncertainties for model-data-mismatch as well as for prior fluxes in both CSR and LUMIA. The spatial flux corrections (innovation of fluxes) shown in Fig. 8 denote quite good agreement between CSR and LUMIA estimates. In this experiment, the differences in June and December decreased to 0.233 and 0.035 PgC $yr^{-1}$, respectively, in comparison with the corresponding differences obtained from the default configurations of both systems. That is to say, the impact of uncertainty definition alone amounts to 0.092 and 0.032 PgC $yr^{-1}$ in June and December, respectively, leading to approximately 30% and 50% of the overall system-to-system differences. The rest of the differences may be attributed to differences in the convergence of the cost function to reach the minimum values.

**How do our results explain the range of uncertainties reported in scientific literature?**

We attempt here to shed light on the spread in our inversions in line with the spread calculated over different ensembles of inversions reported in previous studies by Monteil et al. (2020) in the EUROCOM experiment and by Thompson et al. (2020) dedicated to investigating the impact of 2018 drought on NEE over Europe. Figure 9 shows the spreads amid the three ensembles of: 1) eight inversions conducted in our results denoted as “Ensemble”, 2) six inversions of the EUROCOM experiment “EUROCOM”, and 3) five inversions of the drought study of Thompson et al. (2020) denoted as “Drought”. Note that in EUROCOM and Drought, the tracer inversions differed in the atmospheric regional transport models, the definition of boundary conditions, the definition of control vector, the selection of atmospheric datasets, and the a-priori fluxes. These differences are expected to span a large range of uncertainty sources in the posterior NEE. The climatological monthly estimates of NEE were averaged over “EUROCOM” inversion members for the respective years 2006-2015, except for one inversion (NAME), which was limited to 2011-2015. “Ensemble” and “Drought” were confined to the analysis year of 2018. The monthly NEE estimates were calculated for all ensembles as the average over their respective inversion

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

members. The annual mean of NEE estimated with "EUROCOM", "Ensemble", and "Drought" amounts to -0.19 and -0.29, and -0.05 (PgC) with standard deviations of 0.34 and 0.29, and 0.46 (PgC), respectively.

The spreads amid each ensemble of inversions are illustrated by the min and max values bounded around the mean on the error bars (Fig. 9). The monthly mean of NEE estimates shows a good consistency in all the ensembles. The spreads are also relatively comparable, albeit variable over months. For instance, "EUROCOM" and "Drought" exhibit larger spreads during the growing season (April-August), while "Ensemble" has a larger spread in the rest of months -i.e., during winter. Noteworthy, all ensembles experience large spreads during June and May. Although the participating inversions to "EUROCOM" and "Drought" had different configurations, the spreads were not largely different from our inversion spreads. This implies that atmospheric transport is likely responsible for the large differences in posterior fluxes. Our results also suggest that the definition of uncertainty covariance matrices and lateral boundary conditions lead to a non-negligible impact. Moreover, the discrepancies in "EUROCOM" and "Drought" estimates are expected to be partially caused by using different atmospheric datasets in the inversion systems. Munassar et al. (2022) found that posterior fluxes can be more sensitive to changing the number of stations than changing the prior flux models.

## 5 Conclusions

Estimating atmospheric tracer fluxes through inverse modeling systems has been widely used, in particular targeting the major GHGs to improve the quantification of natural (both terrestrial and oceanic) sources and sinks. Here, an analysis of differences in posterior fluxes of $CO_2$ was carried out using inversion systems deploying different regional transport models to assess the impact of atmospheric transport on inverse modeling estimates. The difference between minimum spread and maximum spread for annually integrated fluxes was found to be 0.92 PgC $yr^{-1}$ for the ensemble range of 0.20 and -0.72 PgC $yr^{-1}$ with a mean estimate of -0.29 PgC $yr^{-1}$ calculated over the full domain of Europe in 2018. The standard deviation of the ensemble members amounted to 0.30 PgC $yr^{-1}$. We found that transport model uncertainties lead to large spatial and temporal discrepancies in the optimized fluxes. Temporal and spatial differences of posterior fluxes are consistent with the differences in simulated $CO_2$ concentration sampled with STILT and FLEXPART over the station network. However, they demonstrate a spatial pattern over certain areas during June and December suggesting rather systematic differences between STILT and FLEXPART. Furthermore, the variable magnitude of differences in prior concentrations also reflect the random errors of models to represent the atmospheric mesoscale motion over the receptor locations. The findings outline that the largest part of the differences is predominantly caused by differences in the estimation of $CO_2$ mixing by the transport models. Meteorological forcing data partially contribute to the differences, especially during the months in which net release of $CO_2$ occurs. However, the differences in $CO_2$ simulations did not show large sensitivity to other parameters such as the way the surface layer height (maximum altitude considered sensitive to the fluxes in Lagrangian models) and the mixing height are defined. In addition, the global transport models used to provide the far field contributions to the regional domain

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

are responsible for small, but non-negligible differences in the inversion estimates. These differences appeared to be homogeneous spatially and temporally, which can be considered as bias-like. The differences arising from using different inversion systems integrated over the entire domain of Europe were on the contrary rather small, once differences such as the transport model and the uncertainties are controlled for. However, such an impact is partially a result of applying different structure and shape in prior flux uncertainty, while the rest may be attributed to differences in the cost function convergence to reach the minimum. This reflects the importance of the way the uncertainty is prescribed in the tracer inversion systems. The large spread resulting from swapping the regional transport model emphasizes the need for further evaluation of atmospheric transport models to improve the performance of the models. At the same time, it is important to realistically account for the transport errors in the tracer inversions. Errors in meteorology parameters assimilated in transport models as forcing data should also be accounted for explicitly, potentially through making use of an ensemble of meteorology data to estimate such errors. Despite the non-negligible difference between inversion systems, this study indicates the importance of following a common inversion protocol when reporting flux estimates from different inversion frameworks.

**Code and data availability**

The simulations of the ensemble of inversions (posterior NEE calculated using CSR and LUNIA), their respective prior fluxes, and codes can be made available upon request to the corresponding author. The atmospheric datasets of $CO_2$ dry mole fractions are available at the ICOS Carbon Portal and can be accessed from https://doi.org/10.18160/ERE9-9D85 (Drought 2018 Team and ICOS Atmosphere Thematic Centre, 2020).

**Competing interests**

At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics. The peer-review process was guided by an independent editor, and the authors also have no other competing interests to declare.

**Acknowledgements**

The authors thank Mathias Göckede for his valuable comments on the manuscript in the internal review. SM, CG, CR, and F-T K, acknowledge the computational support of Deutsches Klimarechenzentrum (DKRZ) where the CSR inversion system is implemented. The authors acknowledge the use of the atmospheric dataset of $CO_2$ dry mole fractions collected throughout ICOS and NOAA site network.

**Financial support**

This research has been supported by Horizon 2020 (VERIFY (grant no. 776810)).

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

## 6 References

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual atmospheric $CO_2$ sources and sinks: 1. Method and control inversion, Journal of Geophysical Research: Atmospheres, 104, 26161-26178, https://doi.org/10.1029/1999JD900342, 1999.

Chen, H. W., Zhang, F. Q., Lauvaux, T., Davis, K. J., Feng, S., Butler, M. P., and Alley, R. B.: Characterization of Regional-Scale $CO_2$ Transport Uncertainties in an Ensemble with Flow-Dependent Transport Errors, Geophysical Research Letters, 46, 4049-4058, 10.1029/2018gl081341, 2019.

Deng, A. J., Lauvaux, T., Davis, K. J., Gaudet, B. J., Miles, N., Richardson, S. J., Wu, K., Sarmiento, D. P., Hardesty, R. M., Bonin, T. A., Brewer, W. A., and Gurney, K. R.: Toward reduced transport errors in a high resolution urban $CO_2$ inversion system, Elementa-Sci Anthrop, 5, ARTN 20
10.1525/elementa.133, 2017.

Engelen, R. J.: On error estimation in atmospheric $CO_2$inversions, Journal of Geophysical Research, 107, 10.1029/2002jd002195, 2002.

Enting, I. G., and Newsam, G. N.: Inverse Problems in Atmospheric Constituent Studies .2. Sources in the Free Atmosphere, Inverse Probl, 6, 349-362, Doi 10.1088/0266-5611/6/3/005, 1990.

Fletcher, S. E. M., Gruber, N., Jacobson, A. R., Gloor, M., Doney, S. C., Dutkiewicz, S., Gerber, M., Follows, M., Joos, F., Lindsay, K., Menemenlis, D., Mouchet, A., Muller, S. A., and Sarmiento, J. L.: Inverse estimates of the oceanic sources and sinks of natural $CO_2$ and the implied oceanic carbon transport, Global Biogeochemical Cycles, 21, Artn Gb1010
10.1029/2006gb002751, 2007.

Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A. T., Dargaville, R., Aalto, T., Brandt, J., Christensen, J. H., Frohn, L. M., Haszpra, L., Karstens, U., Rodenbeck, C., Ramonet, M., Carboni, G., and Santaguida, R.: Comparing atmospheric transport models for future regional inversions over Europe - Part 1: mapping the atmospheric $CO_2$ signals, Atmospheric Chemistry and Physics, 7, 3461-3479, DOI 10.5194/acp-7-3461-2007, 2007.

Gerbig, C., Korner, S., and Lin, J. C.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, Atmospheric Chemistry and Physics, 8, 591-602, DOI 10.5194/acp-8-591-2008, 2008.

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, Journal of Geophysical Research: Atmospheres, 106, 4703-4715, 2001.

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

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.: Technical Note: Atmospheric $CO_2$ inversions on the mesoscale using data-driven prior uncertainties: methodology and system evaluation, Atmos. Chem. Phys., 18, 3027-3045, 10.5194/acp-18-3027-2018, 2018a.

Kountouris, P., Gerbig, C., Rödenbeck, C., Karstens, U., Koch, T. F., and Heimann, M.: Atmospheric $CO_2$ inversions on the mesoscale using data-driven prior uncertainties: quantification of the European terrestrial $CO_2$ fluxes, Atmos. Chem. Phys., 18, 3047-3064, 10.5194/acp-18-3047-2018, 2018b.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J., and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, J Geophys Res-Atmos, 108, Artn 4493
10.1029/2002jd003161, 2003.

Liu, J. J., Fung, I., Kalnay, E., and Kang, J. S.: $CO_2$ transport uncertainties from the uncertainties in meteorological fields, Geophysical Research Letters, 38, Artn L12808
10.1029/2011gl047213, 2011.

Mahadevan, P., Wofsy, S. C., Matross, D. M., Xiao, X. M., Dunn, A. L., Lin, J. C., Gerbig, C., Munger, J. W., Chow, V. Y., and Gottlieb, E. W.: A satellite-based biosphere parameterization for net ecosystem $CO_2$ exchange: Vegetation Photosynthesis and Respiration Model (VPRM), Global Biogeochemical Cycles, 22, Artn Gb2005
10.1029/2006gb002735, 2008.

Monteil, G., Broquet, G., Scholze, M., Lang, M., Karstens, U., Gerbig, C., Koch, F. T., Smith, N. E., Thompson, R. L., Luijkx, I. T., White, E., Meesters, A., Ciais, P., Ganesan, A. L., Manning, A., Mischurow, M., Peters, W., Peylin, P., Tarniewicz, J., Rigby, M., Rödenbeck, C., Vermeulen, A., and Walton, E. M.: The regional European atmospheric transport inversion comparison, EUROCOM: first results on European-wide terrestrial carbon fluxes for the period 2006–2015, Atmos. Chem. Phys., 20, 12063-12091, 10.5194/acp-20-12063-2020, 2020.

Monteil, G., and Scholze, M.: Regional $CO_2$ inversions with LUMIA, the Lund University Modular Inversion Algorithm, v1.0, Geosci. Model Dev., 14, 3383-3406, 10.5194/gmd-14-3383-2021, 2021.

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

Munassar, S., Rödenbeck, C., Koch, F. T., Totsche, K. U., Gałkowski, M., Walther, S., and Gerbig, C.: Net ecosystem exchange (NEE) estimates 2006–2019 over Europe from a pre-operational ensemble-inversion system, Atmos. Chem. Phys., 22, 7875-7892, 10.5194/acp-22-7875-2022, 2022.

Peylin, P., Baker, D., Sarmiento, J., Ciais, P., and Bousquet, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric $CO_2$ data, J Geophys Res-Atmos, 107, Artn 4385 10.1029/2001jd000857, 2002.

Pisso, I., Sollum, E., Grythe, H., Kristiansen, N. I., 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., Seibert, P., and Stohl, A.: The Lagrangian particle dispersion model FLEXPART version 10.4, Geosci. Model Dev., 12, 4955-4997, 10.5194/gmd-12-4955-2019, 2019.

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

Rödenbeck, C.: Estimating $CO_2$ sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport, 06, 2005.

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

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

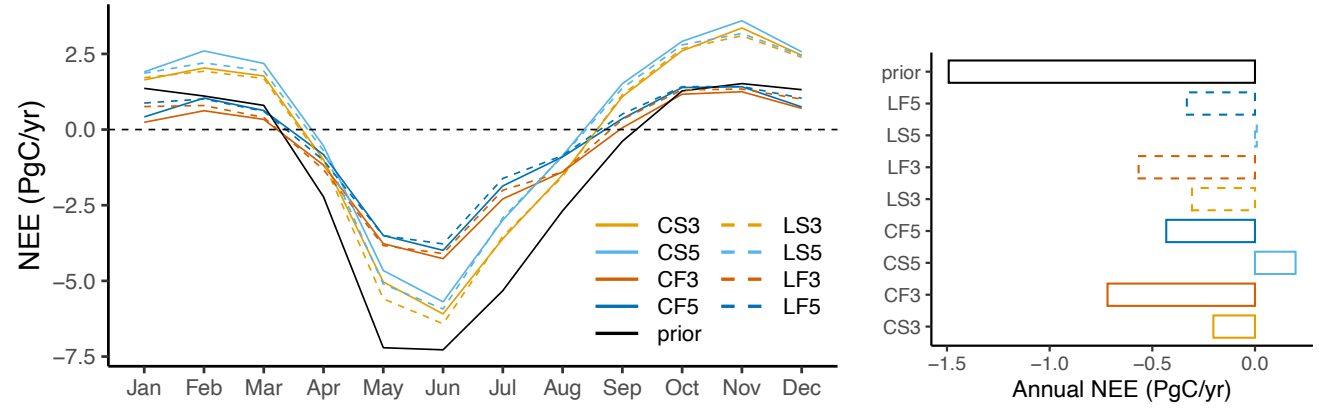
Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., Chevallier, F., Crowell, S., Davis, K. J., Deng, F., Denning, S., Feng, L., Jones, D., Liu, J. J., and Palmer, P. I.: Quantifying the Impact of Atmospheric Transport Uncertainty on $CO_2$ Surface Flux Estimates, Global Biogeochemical Cycles, 33, 484-500, 10.1029/2018gb006086, 2019.

Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., and Mukai, H.: The CO<sub>2</sub> release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, Atmos. Chem. Phys., 11, 6855-6870, 10.5194/acp-11-6855-2011, 2011.

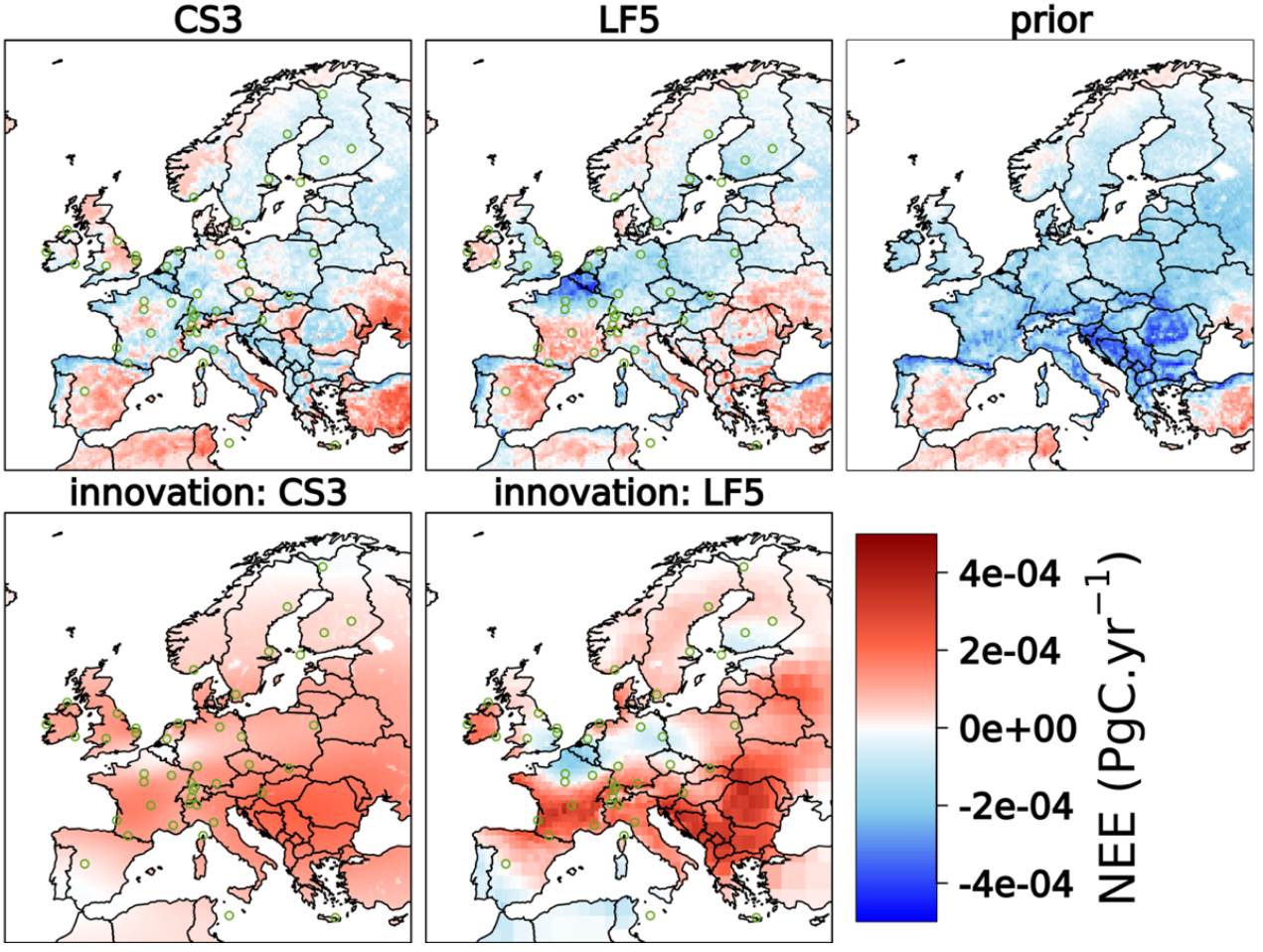
Thompson, R. L., Broquet, G., Gerbig, C., Koch, T., Lang, M., Monteil, G., Munassar, S., Nickless, A., Scholze, M., Ramonet, M., Karstens, U., van Schaik, E., Wu, Z., and Rodenbeck, C.: Changes in net ecosystem exchange over Europe during the 2018 drought based on atmospheric observations, Philos Trans R Soc Lond B Biol Sci, 375, 20190512, 10.1098/rstb.2019.0512, 2020.

Tolk, L. F., Meesters, A. G. C. A., Dolman, A. J., and Peters, W.: Modelling representation errors of atmospheric $CO_2$ mixing ratios at a regional scale, Atmospheric Chemistry and Physics, 8, 6587-6596, DOI 10.5194/acp-8-6587-2008, 2008.

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


595 **Figure 1: Left panel refers to posterior monthly NEE estimated using eight inversions, including prior NEE shown in black colour, with CSR (solid lines) and LUMIA (dashed lines), and right panel denotes the corresponding annually aggregated fluxes. Orange and red colours correspond to TM3 and dark/light blue to TM5. Orange and light blue colours refer to STILT and red and dark blue to FLEXPART.**



https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

**Figure 2: First row shows the spatial distributions of annual NEE estimated with the base inversions CS3 and LF5, as well as their prior. Second row depicts the innovations of fluxes calculated for the inversions CS3, LF5. Green circles denote the locations of observational sites.**

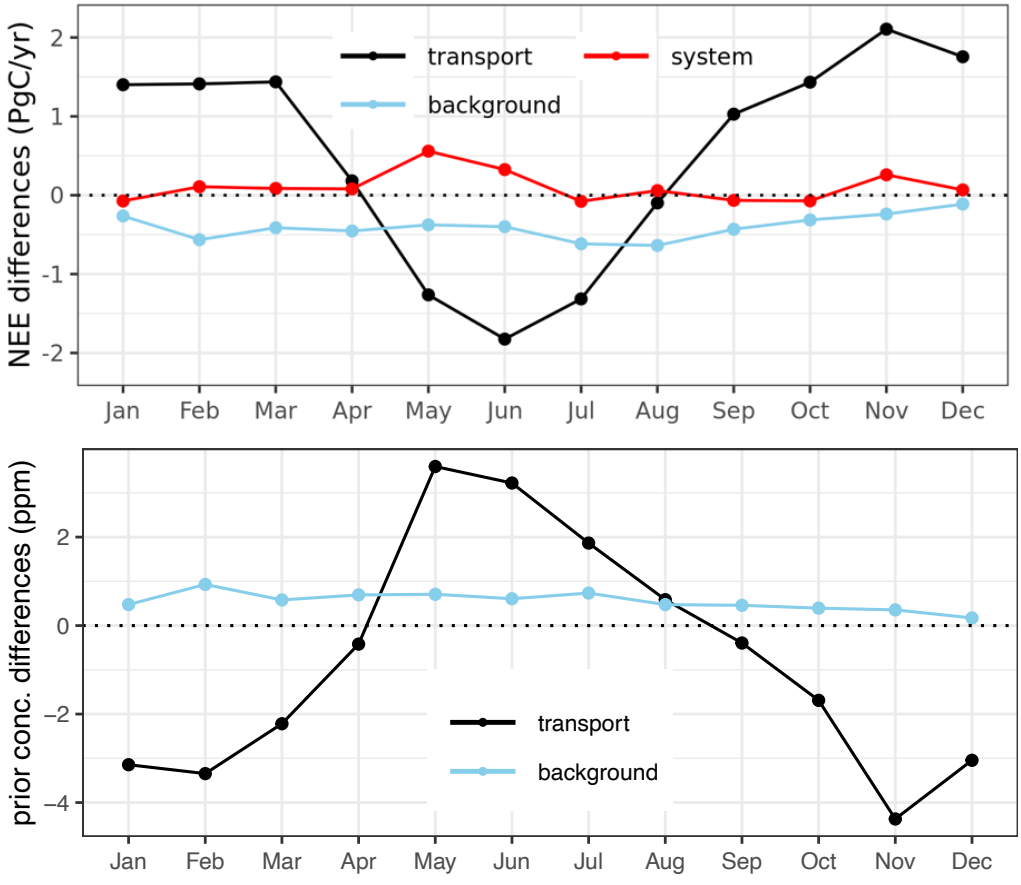



**Figure 3: Differences in optimized fluxes (top) and prior concentrations (bottom) calculated with the regional transport models STILT and FLEXPART (CS3-CF3) and background provided through TM3 and TM5 (CS3-CS5). "system" refers to the differences between CSR and LUMIA inversion for optimized fluxes (CS5-LS5).**

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022

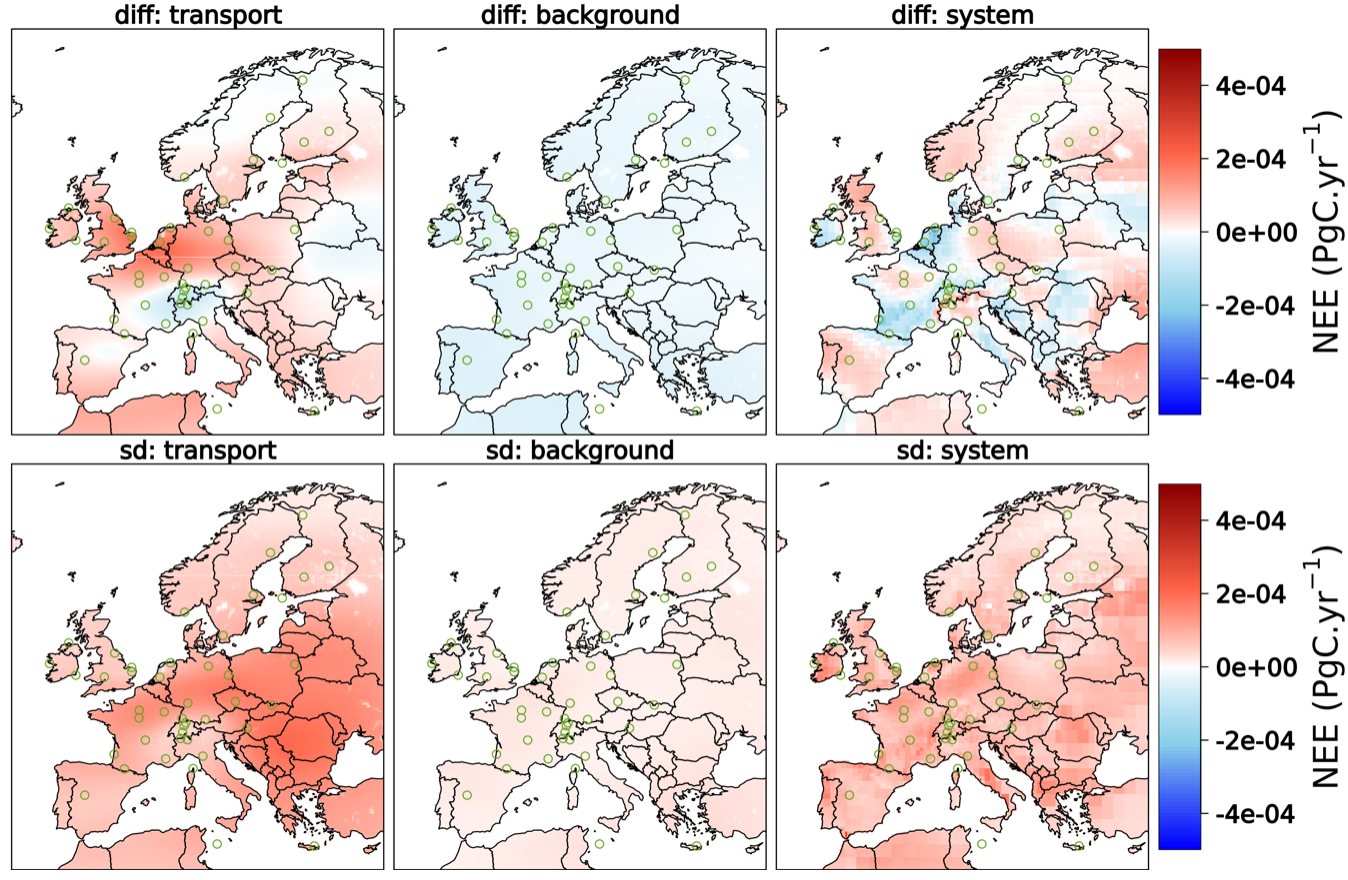
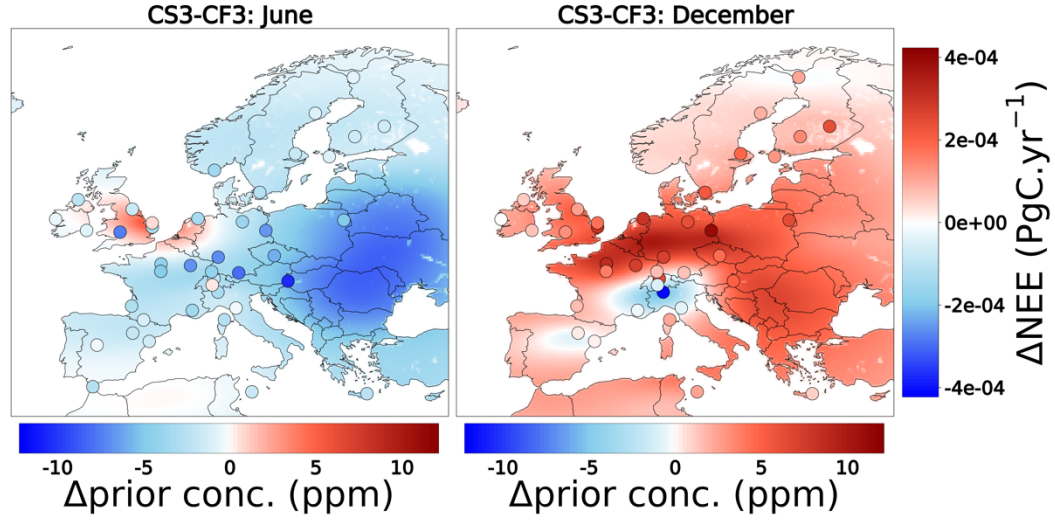




**Figure 4: First row indicates differences in annual posterior NEE estimated with STILT and FLEXPART models referred to as "transport" (CS3-CF3), TM3 and TM5 referred to as "background" (CS3-CS5), and CSR and LUMIA referred to as "system" (CF3-LF3); second row demonstrates the standard deviations of the corresponding monthly differences.**



https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

**Figure 5: Spatial differences of posterior NEE estimated from the inversions CS3 and CF3 with STILT and FLEXPART transport models during June and December; filled circles indicate the differences in prior concentrations at the locations of sites (horizontal legend explains the magnitude of differences).**

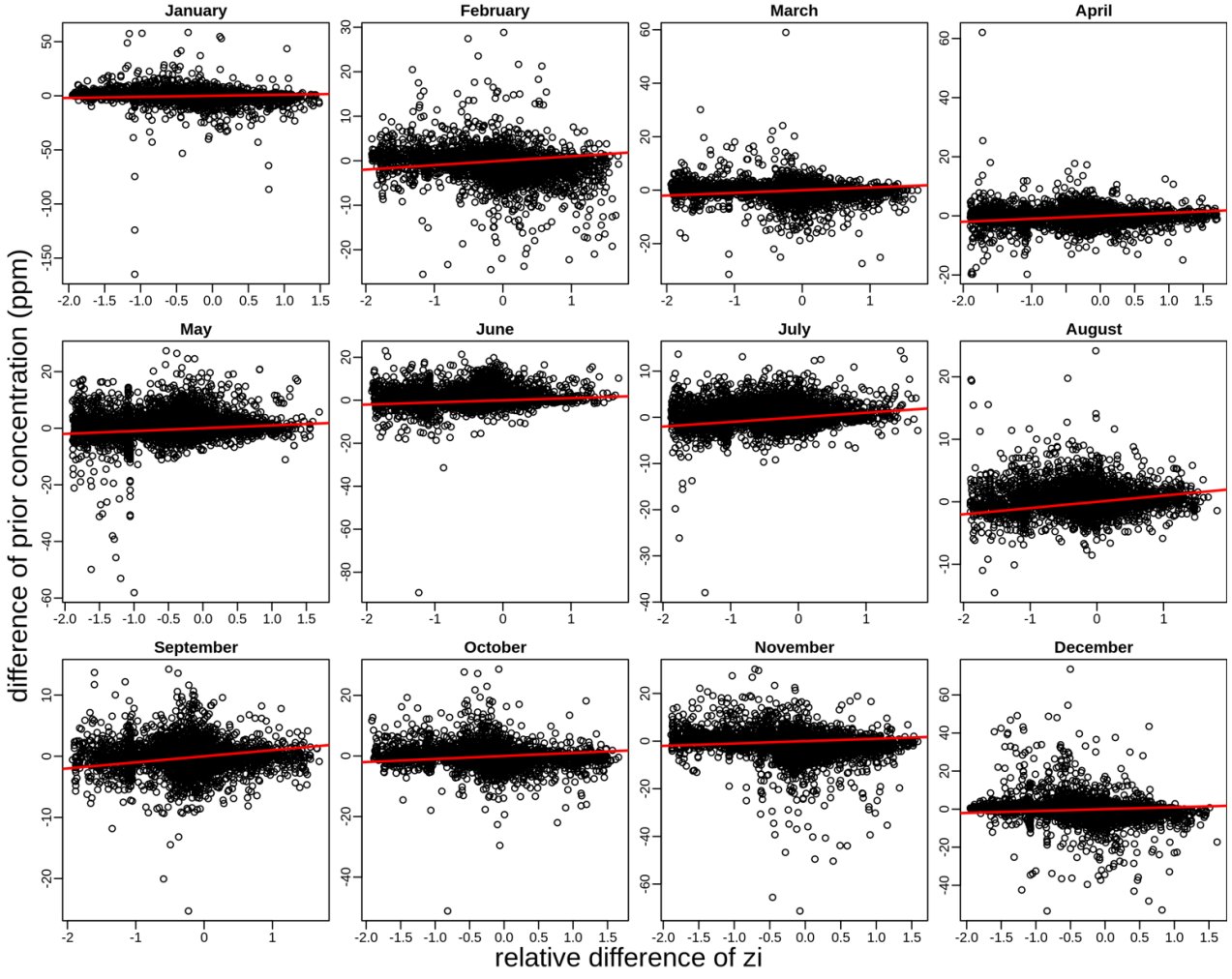



**Figure 6: Scatter plot of differences of prior concentrations and mixing heights calculated with STILT and FLEXPART models (i.e., STILT-FLEXPART on the x- and the y-axis). Red lines indicate the slopes.**

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022

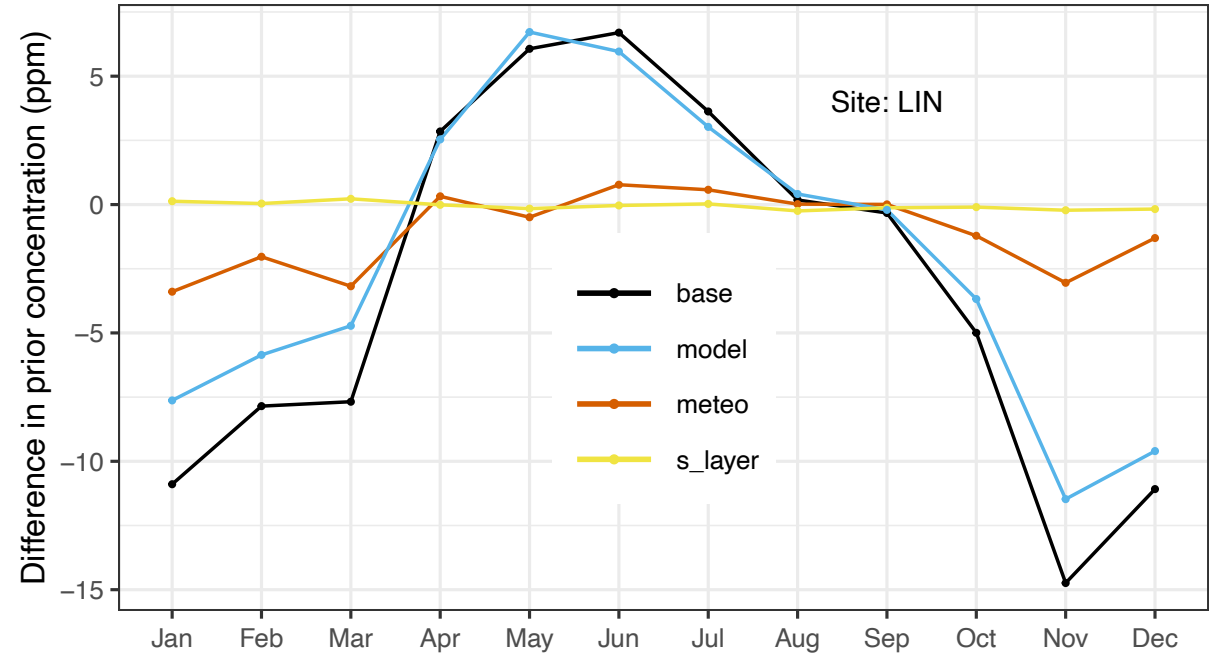




**Figure 7: Differences in prior concentration simulated at LIN with STILT and FLEXPART using different configurations. "s_layer", yellow line, refers to the difference calculated with STILT using two assumptions of defining the surface layer height, once with the default as 0.5 of the mixed layer, and once with 100 m as used in FLEXPART. "meteo", red line, indicates the differences calculated with FLEXPART using two different types of meteorological data, IFS (the STILT default) and ERA-5. "model", blue line, denotes the differences calculated with STILT and FLEXPART, given identical meteorological data (IFS) and surface layer height (100 m). "base", black line, refers to the base configurations of STILT and FLEXPART encompassing all possible differences between models - i.e., 1) STILT with IFS forecasting data and a surface layer height as 0.5 of the mixed layer height, and 2) FLEXPART with ERA-5 reanalysis and the surface layer height of 100 m.**

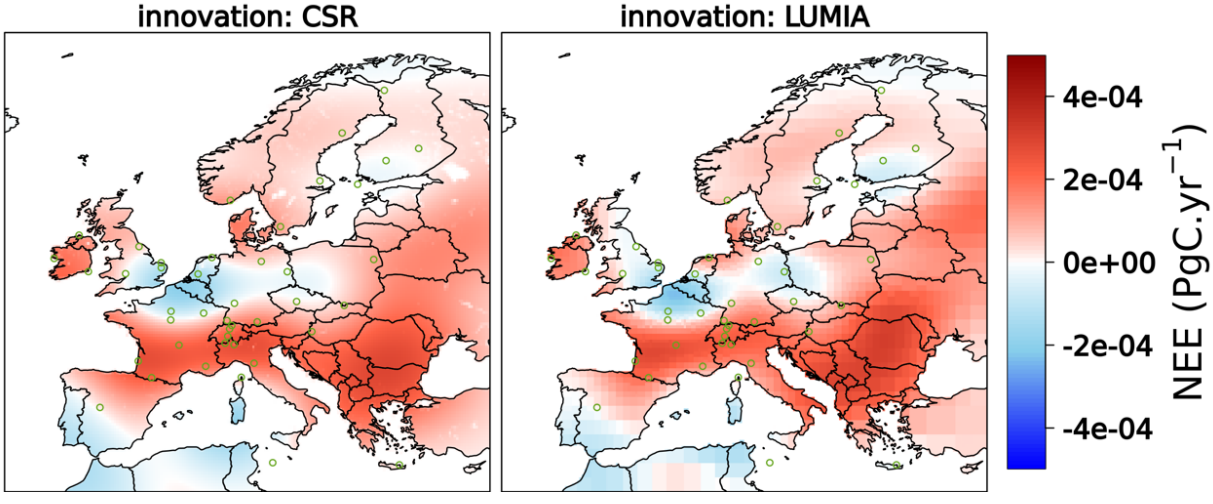
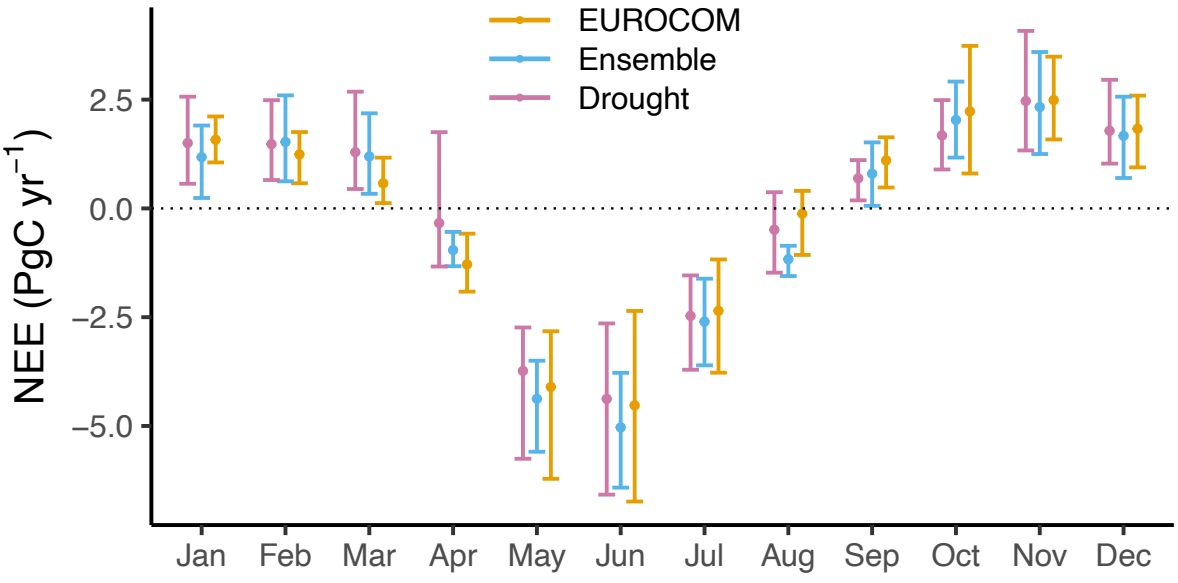



**Figure 8: Innovation of fluxes calculated from CSR and LUMIA using identical uncertainties of prior flux and measurements. The uncertainty flux shape was flat and the decaying spatial correlation was fit to Gaussian function with 500 km scale. FLEXPART and TM5 models were used in this experiment.**

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions




635 **Figure 9: Comparison of monthly NEE estimates calculated as the mean of six inversions taken from Monteil et al. (2020), denoted as "EUROCOM", eight inversion members conducted in our study (set-ups listed in Table 2), denoted as "Ensemble", and five inversions used in Thompson et al. (2020) for the 2018 drought study denoted as "Drought". The error bars refer to the spreads (min/max) over the respective members amid each ensemble of inversions.**

640

645

650

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

655

**Table 1. Atmospheric sites used in the inversions.**

| Site code | Site name | Coordinates (lat, lon)° | STILT release height (magl) | FLEXPART release height (magl) | Time window (UTC) | Uncertainty (ppm) |
|---|---|---|---|---|---|---|
| SM3 | Hyytiala | 61.85, 24.29 | 125 | 125 | 10:00-14:00 | 1.5 |
| BI5 | Bialystok | 53.23, 23.03 | 300 | 300 | 10:00-14:00 | 1.5 |
| FKL | Finokalia | 35.34, 25.67 | 15 | 15 | 10:00-14:00 | 1.5 |
| PAL | Pallas | 67.97, 24.12 | 12 | 12 | 10:00-14:00 | 2.5 |
| PUI | Puijo | 62.91, 27.65 | 84 | 84 | 10:00-14:00 | 1.5 |
| UTO | Uto Baltic Sea | 59.78, 21.37 | 57 | 57 | 10:00-14:00 | 1.5 |
| BIR | Birkenes Observatory | 58.389, 8.25 | 3 | 3 | 11:00-15:00 | 2.5 |
| BR5 | Beromuenster | 47.19, 8.17 | 212 | 212 | 11:00-15:00 | 1.5 |
| DEC | Deltade_lEbre | 40.74, 0.79 | 10 | 10 | 11:00-15:00 | 1.5 |
| EEC | El_Estrecho | 36.0586, -5.664 | 20 | 20 | 11:00-15:00 | 1.5 |
| GIC | Sierra de Gredos | 40.3457, -5.1755 | 20 | 20 | 11:00-15:00 | 2.5 |
| HEI | Heidelberg | 49.417, 8.674 | 30 | 30 | 11:00-15:00 | 4 |
| HP4 | Hohenpeissenberg | 47.8011, 11.0246 | 300 | 131 | 11:00-15:00 | 1.5 |
| ER2 | ERSA | 42.9692, 9.3801 | 40 | 40 | 11:00-15:00 | 1.5 |
| HT3 | Hyltemossa | 56.0969, 13.4189 | 150 | 150 | 11:00-15:00 | 1.5 |
| HU4 | Hegyhatsal Tower | 46.95, 16.65 | 115 | 115 | 11:00-15:00 | 1.5 |
| IP3 | Ispra | 45.8147, 8.636 | 100 | 100 | 11:00-15:00 | 1.5 |
| KR3 | Kresin | 49.572, 15.08 | 250 | 250 | 11:00-15:00 | 1.5 |
| LMU | La Muela | 41.5941, -1.1003 | 80 | 79 | 11:00-15:00 | 1.5 |
| LMP | Lampedusa | 35.53, 12.62 | 10 | 10 | 11:00-15:00 | 1.5 |
| LUT | Lutjewad | 53.4036, 6.3528 | 60 | 60 | 11:00-15:00 | 2.5 |
| NO3 | Norunda | 60.0864, 17.4794 | 100 | 100 | 11:00-15:00 | 1.5 |
| SV3 | Svartberget | 64.256, 19.775 | 150 | 150 | 11:00-15:00 | 1.5 |
| TR4 | Trainou | 47.9647, 2.1125 | 180 | 180 | 11:00-15:00 | 1.5 |
| OHP | Obervatoire de Haute Provence | 43.931, 5.712 | 100 | 100 | 11:00-15:00 | 1.5 |
| SA3 | Saclay | 48.7227, 2.142 | 100 | 100 | 11:00-15:00 | 1.5 |
| LHW | Laegern Hochwacht | 47.4822, 8.3973 | 400 | 32 | 11:00-15:00 | 2.5 |
| BS3 | Bilsdale | 54.359, -1.15 | 248 | 248 | 12:00 -16:00 | 1.5 |
| RG2 | Ridge Hill | 51.9976, -2.54 | 90 | 90 | 12:00 -16:00 | 1.5 |
| TA3 | Tacnolestan | 52.5177, 1.1386 | 185 | 185 | 12:00 -16:00 | 1.5 |
| WAO | Weybourne Norfolk | 52.9502, 1.1219 | 10 | 10 | 12:00 -16:00 | 1.5 |
| OP3 | OPE_ANDRA | 48.5619, 5.5036 | 120 | 120 | 14:00-17:00 | 1.5 |
| GA5 | Gartow | 53.0657, 11.4429 | 341 | 341 | 14:00-18:00 | 1.5 |
| LIN | Lindenberg | 52.1663, 14.1226 | 98 | 98 | 14:00-18:00 | 1.5 |
| BIS | Biscarrose | 44.3781, -1.2311 | 47 | 47 | 14:00-18:00 | 2.5 |
| CRP | Carnoise Point | 52.18, -6.37 | 14 | 14 | 14:00-18:00 | 1.5 |
| MHD | MaceHead | 53.3261, -9.9036 | 24 | 24 | 14:00-18:00 | 1.5 |
| MLH | Marlin Head | 55.355, -7.333 | 47 | 47 | 14:00-18:00 | 1.5 |
| JFJ | Jungfraujoch | 46.5475, 7.9851 | 720 | 3570 | 23:00-3:00 | 1.5 |
| KAS | Kasrprovy Wierch | 49.2325, 19.9818 | 480 | 1989 | 23:00-3:00 | 1.5 |
| PUY | Puy de Dome | 45.7719, 2.9658 | 400 | 1465 | 23:00-3:00 | 1.5 |

https://doi.org/10.5194/acp-2022-510
Preprint. Discussion started: 1 August 2022


Atmospheric Chemistry and Physics Discussions

| | | | | | | |
|---|---|---|---|---|---|---|
| SI2 | Schauinsland | 47.91, 7.91 | 450 | 1205 | 23:00-3:00 | 1.5 |
| PTR | Plateau Rosa Station | 45.94, 7.71 | 500 | 3480 | 23:00-3:00 | 1.5 |
| PD2 | Pic du Midi | 42.9372, 0.1411 | 1458 | 2877 | 23:00-3:00 | 1.5 |
| CMN | Monte Cimone | 44.1963, 10.6999 | 670 | 2165 | 23:00-3:00 | 1.5 |

**Table 2: List of the inversion set-ups**

| **Inversion system** | **Transport model** | **Global boundary condition** | **Identifier code** | **Flux Uncertainty** | |
|---|---|---|---|---|---|
| | | | | Shape | Decay |
| **LUMIA** | FLEXPART | TM5 | LF5 | Variable | Gaussian |
| **LUMIA** | FLEXPART | TM3 | LF3 | Variable | Gaussian |
| **LUMIA** | STILT | TM5 | LS5 | Variable | Gaussian |
| **LUMIA** | STILT | TM3 | LS3 | Variable | Gaussian |
| **CSR** | STILT | TM3 | CS3 | Flat | Hyperbolic |
| **CSR** | STILT | TM5 | CS5 | Flat | Hyperbolic |
| **CSR** | FLEXPART | TM3 | CF3 | Flat | Hyperbolic |
| **CSR** | FLEXPART | TM5 | CF5 | Flat | Hyperbolic |