**Atmospheric
Chemistry
and Physics**

# A Bayesian inversion estimate of $N_2O$ emissions for western and central Europe and the assessment of aggregation errors

**R. L. Thompson[1,2], C. Gerbig[2], and C. Rödenbeck[2]**

[1]Laboratoire des Sciences du Climat et l'Environnement (LSCE), UMR 8212, Paris, France
[2]Max Planck Institute for Biogeochemistry (MPI-BGC), Jena, Germany

**Abstract.** A Bayesian inversion approach was used to retrieve temporally and spatially resolved $N_2O$ fluxes for western and central Europe using in-situ atmospheric observations from the tall tower site at Ochsenkopf, Germany (50°01′ N, 11°48′ E). For atmospheric transport, the STILT (Stochastic Time-Inverted Lagrangian Transport) model was employed, which was driven with ECMWF analysis and short term forecast fields. The influence of temporal aggregation error, as well as the choice of spatial and temporal correlation scale length, on the retrieval was investigated using a synthetic dataset consisting of mixing ratios generated for the Ochsenkopf site. We found that if the aggregation error is ignored, then a significant bias error in the retrieved fluxes ensues. However, by estimating this error and projecting it into the observation space, it was possible to avoid bias errors in the retrieved fluxes. Using the real observations from the Ochsenkopf site, $N_2O$ fluxes were retrieved every 7 days for 2007 at 2 by 2 degrees spatial resolution. Emissions of $N_2O$ were strongest during the summer and autumn months, with peak emissions in August and September, while the regions of Benelux and northern United Kingdom had strongest annual mean emissions.

## 1 Introduction

The nitrous oxide ($N_2O$) mixing ratio has been increasing in the atmosphere since the industrial revolution, from 270 ppb in 1750 (Ramaswamy, 2001) to 319 ppb in 2005 (Forster et al., 2007) with a steady growth rate of around 0.26% since the early 1980's. The increase in $N_2O$ is worrisome for two main reasons. First, it is a greenhouse gas; this means that its atmospheric increase translates to an enhancement in radia-

tive forcing of $0.16\pm0.02\,\mathrm{Wm^{-2}}$ (Ramaswamy, 2001) making it currently the fourth most important long-lived greenhouse gas and is predicted to soon overtake CFC's to become the third most important (Forster et al., 2007). Second, it plays an important role in stratospheric ozone destruction (Crutzen, 1974; Nevison and Holland, 1997) and in the 21$^{st}$ century is the main ozone-depleting substance and is slowing the recovery of the ozone hole (Ravishankara et al., 2009).

Human activities are the primary cause of the atmospheric $N_2O$ increase. The largest anthropogenic source of $N_2O$ is from the use of N-fertilizers in agriculture (Olivier et al., 1998), which greatly enhances soil $N_2O$ emissions via microbial nitrification and denitrification pathways (e.g. Bremner, 1997). Other important anthropogenic sources of $N_2O$ are fossil fuel combustion and industrial processes, such as adipic and nitric acid production (Olivier et al., 1998). There are also important natural sources of $N_2O$, the largest of which is from soils under natural vegetation, but the open ocean, coastal upwelling regions, continental shelves, and estuaries are significant sources too (Kroeze et al., 2005; Nevison et al., 2004). Moreover, it is important to note, that as anthropogenic N permeates the nitrogen cycle, the natural emissions of $N_2O$ are also being enhanced (Galloway et al., 2008; Gruber and Galloway, 2008).

There have been few previous top-down estimates of $N_2O$ fluxes using atmospheric observations. On a regional scale, there have been a small number of $N_2O$ emission estimates based on concurrent measurements of $N_2O$ and radon, where radon is used to determine the dilution of $N_2O$ emissions in the atmosphere, such as Messager et al. (2008), Biraud et al. (2002) and Schmidt et al. (2001), and two estimates based on an inversion of atmospheric $N_2O$ observations (Corazza et al., 2011; Manning et al., 2003). Another modelling study by Kort et al. (2008) evaluated emission inventories by comparison of simulated tracer distributions using an atmospheric transport model with aircraft based measurements of $N_2O$ in North America and inferred a scaling of fluxes that was

not spatially or temporally resolved. On the global scale, the only inversion estimates are those of Huang et al. (2008), Hirsch et al. (2006) and Prinn et al. (1990) and these have been at low resolution. The shortage of top-down estimates is partly due to the relative scarcity of data but also to problems in the calibration scales used for N$_2$O measurements by different laboratories, which has made comparisons between stations difficult.

There have been several recent initiatives that will improve the availability of data as well as make measurements between stations more comparable. In Europe, for example, the establishment of a new network of tall towers under the framework of the CHIOTTO project (Continuous HIgh-precisiOn Tall Tower Observations of greenhouse gases) provides in-situ measurements of N$_2$O from 7 stations across Europe. Furthermore, these stations participate in regular inter-comparisons, ensuring a uniform scale. These in-situ measurements provide N$_2$O observations at approximately hourly frequencies making it possible to resolve short-term variability in the atmosphere. The question, however, is how to best incorporate high-resolution measurements into an inversion scheme.

In an ideal inversion, there would be multiple observations of each state variable, in our case the fluxes, at each resolved time-step. In reality, though, a significant portion of the state variables may be undetected by the observations, owing to the fact that there are only a discrete number of measurement sites and the signal observed at these is dependent on the atmospheric circulation. Furthermore, one must make assumptions about the variability of the true fluxes on time-scales shorter than that of the resolved fluxes. To minimise the interval over which these assumptions are made, one can define a higher temporal resolution for the fluxes but there is a trade-off between increasing the temporal resolution and the number of observations available to constrain the fluxes at each time-step. A similar problem occurs when defining the spatial resolution as one assumes that the fluxes within the resolved area are not significantly different from the mean of the area. Errors resulting when such assumptions are invalid are defined as aggregation errors (Kaminski et al., 2001; Peylin et al., 1999).

In some cases the temporal variability of fluxes over certain time intervals can be described and used in the inversion. For example, for CO$_2$ fluxes, the behaviour on diurnal time-scales can be described and thus the variability on time intervals of less than 1 day does not need to be explicitly solved in the inversion as long as it is taken into account (Peylin et al., 2005). However, in the case of N$_2$O there is no regular pattern to the short-term flux variability, hence the approach used to account for short-term variability in CO$_2$ fluxes is not applicable. Thus the problem remains; on one hand, if high temporal resolution is used, then there are fewer observations available to constrain the fluxes and there are computational limits on how much the resolution may be increased. On the other hand, if lower temporal resolution is used, and if the real short-term variability of the fluxes and their influence on the observations is not accounted for, then a significant risk of temporal aggregation error ensues.

In this paper, we first investigate how high-frequency observations may best be utilized for retrieving N$_2$O fluxes and how sensitive the inversion is to temporal aggregation errors. This is investigated using synthetic fluxes and mixing ratios but with atmospheric transport determined from meteorological reanalysis data. Secondly, we investigate the potential to use in-situ measurements from the observation site, Ochsenkopf in Germany (50°01′ N, 11°48′ E), to constrain N$_2$O emissions and present new estimates. The tall tower observatory is located on the summit of the mountain, Ochsenkopf, in the Fichtelgebirge range, at an elevation of 1022 m. The Fichtelgebirge region is extensively forested and has a low population density (for further details about the site refer to Thompson et al. (2009). It has been shown that air sampled from the highest level on the tower (163 m a.g.l.), represents the well mixed boundary layer and has a footprint extending over a large part of western and central Europe (Thompson et al., 2009).

The paper is structured as follows. In Sect. 2, we describe the measurements, the atmospheric transport model, STILT, used to describe the relationship between the observations and the fluxes, and provide details on the Bayesian inversion scheme used. The synthetic data, which are used to test the temporal aggregation error, are also described in this section. Section 3.1 covers the results of various tests for the influence of temporal resolution, observational data frequency, and the impact that changes to the prior flux error covariance matrix have on the retrieved fluxes. Finally, in Sect. 3.2, the results of the inversion using the real Ochsenkopf observations are presented and discussed.

## 2 Methodology

### 2.1 Measurements

Ochsenkopf has a continuous record of N$_2$O since mid-2006. N$_2$O is measured using a Gas Chromatograph (GC), Agilent 6890, equipped with an electron capture detector. The method was largely developed by Worthy et al. (2003) with further improvements by Popa (2008) (details specific to the Ochsenkopf GC set-up can be found in Thompson et al., 2009). Air measurements are made every 16 minutes and are interspersed with reference measurements. The long-term precision is 0.18 ppb (this is the mean over 2.5 years of the standard deviation of 3 consecutive measurements of a gas standard that is measured daily) and the accuracy is 0.5 ppb. Air is sampled from 3 heights on the tall tower, 23, 90 and 163 m, which are measured on a 3-hourly cycle (note that for the inversion in this paper, we use only the data from the 163 m level).

## 2.2 Atmospheric transport model

The STILT (Stochastic Time-Inverted Lagrangian Transport) model is used to simulate the transport of air parcels backwards in time from a receptor point, in this case the tall tower, Ochsenkopf (50°01′ N, 11°48′ E). STILT calculates back trajectories from the receptor using wind fields while releasing ensembles of particles, which represent the air parcel. Analysed wind fields are interpolated to the location of each particle and the particles themselves are subject to stochastic perturbations that are parameterised to represent the effects of turbulent transport. The density of particles is used to calculate the surface influence and the footprint (Lin et al., 2003; Lin et al., 2004). Here we define the footprint as the mean surface influence function over a given time interval defined with reference to a specific receptor, e.g. a tall tower. This means that in the backward-time run, emissions upstream of the receptor at a location with higher particle density have a greater contribution to changes in the mixing ratio at the receptor. STILT has already been implemented in a number of regional-scale trace gas experiments (Gerbig et al., 2003a; Gerbig et al., 2003b; Lin et al., 2004; Lin et al., 2007; Macatangay et al., 2008; Matross et al., 2006; Miller et al., 2008).

STILT was driven with meteorological data from ECMWF (0:00 and 12:00 analysis fields combined with short-term forecast fields), which have a temporal resolution of 3-hours and a spatial resolution of $1/4$ by $1/4$ degrees. The STILT model provides surface influence functions with a dynamic resolution, with the finest grid being $1/12$ degrees latitude by $1/8$ degrees longitude, which increases with distance from the receptor to reduce computational time as well as to prevent under sampling of surface fluxes at times when particles are distributed over large areas (Gerbig et al., 2003a). The model domain was chosen to cover Europe, extending from 35° N to 62° N and 168° W to 35° E with the Ochsenkopf site approximately in the centre.

## 2.3 Simulation of N2O

The mixing ratio at the receptor located at $x_r$, $y_r$, $z_r$ at time $t_r$ is dependent on the surface influence function, calculated by STILT, and the spatial and temporal distribution of the tracer emissions. The surface influence $F$ links the surface emissions to concentration changes at the receptor and is determined by the total time the particle ensemble spends in a column of air with height $h$ at $(x_i, y_j, t_m)$, normalized by the total number of particles, and has the dimension of concentration divided by flux (i.e. ppb/µmol m$^{-2}$.s$^{-1}$) (Lin et al., 2003). The column height $h$ represents the height up to which surface fluxes are mixed at each time-step of the STILT model. Gerbig et al. (2003) found that varying $h$ between 10% and 100% of the mixing height, calculated using the modified Richardson number method, did not significantly affect model results, so in this study $h$ was chosen as

one half of the mixing height. The mixing ratio at the receptor $C(x_r, y_r, z_r, t_r)$ is the product of the surface influence and the tracer flux at each time step, summed up over the course of the back-trajectory. This is shown in the equation from Gerbig et al. (2006), which is based on the equation of Uliasz et Pielke (1990):

$$c(x_r, y_r, z_r, t_r) = \sum_{i,j,m} F(x_r, y_r, z_r, t_r | x_i, y_j, t_m) \cdot f(x_i, y_j, t_m)$$
$$+ \sum_{i,j,k} c(x_i, y_j, z_k, t_0) \tag{1}$$

where $F(x_r, y_r, z_r, t_r | x_i, y_j, t_m)$ is the surface influence and relates the flux, $f(x_i, y_i, t_m)$ to the mixing ratio, $c(x_r, y_r, z_r, t_r)$. The last term on the right hand side represents the influence from the boundaries, $c(x_i, y_j, z_k, t_0)$, that is, the contribution to N2O transported from the domain boundaries to the measurement site (where $t_0$ is defined as the time when the back-trajectory exits the domain). The mixing ratios outside the domain are equal to the background amount in the atmosphere, which is assumed not to be directly affected by local and regional sources (the estimate of these mixing ratios is described in more detail below).

The back-trajectories were calculated for 10 days going backwards in time. However, only the influence from particles that remained inside the domain was used and for trajectories that left the domain and re-entered it again, only the influence from particles before they left the domain in the first instance was used.

The contribution to the mixing ratios from the lateral boundary can be estimated by coupling the regional model to a global one, or alternatively, by filtering the observations themselves so that what is left are only the instances when they are largely representative of the free-troposphere or "background" mixing ratio (Wang and Barrett, 2003). For N2O, however, the global models available at the time of this study were either too coarse and/or too uncertain to be used for this purpose and from the observed time-series at Ochsenkopf it is difficult to distinguish periods when the mixing ratio represents the "background", as there is variability occurring on both short and long timescales.

We have, therefore, chosen to use a mixing ratio field derived from the interpolation of the NOAA/ESRL flask data (E. Dlugokencky, personal communication, 2008). For this field, we have only used stations at coastal or high mountain locations to avoid including mixing ratios that have a strong influence from the near-field. In total, 23 stations from the northern hemisphere and tropics were used (see Table 1). Because of missing data and sometimes irregular sampling, all the available data within a 2-week period were averaged for each site. In some cases, only 1 data point was available over the 2-week period and in some instances no data were available. To overcome the problem of missing data, we used a Singular Systems Analysis (SSA) method for gap-filling (Mahecha et al., 2007). In SSA, the frequency components (also known as principle components)

**Table 1.** Northern hemisphere stations used in the interpolation for the background mixing ratio.

| Station | Coordinates | Station | Coordinates |
|---|---|---|---|
| Alert, Canada, ALT | 82.5° N, 62.5° W, 200 m | Mace Head, Ireland, MHD | 53.3° N, 9.9° W, 25 m |
| Ascension Isl., ASC | 7.5° S, 14.4° W, 54 m | Sand Isl., USA, MID | 28.2° N, 177.4° W, 3.7 m |
| Terceira, Azores, AZR | 38.8° N, 27.4° W, 40 m | Mauna Loa, Hawaii, MLO | 19.5° N, 155.6° W, 3397 m |
| Baltic Sea, Poland, BAL | 55.4° N, 17.2° E, 3 m | Niwot Ridge, USA, NWR | 40.0° N, 105.6° W, 3050 m |
| Barrow, Alaska, BRW | 71.3° N, 156.6° W, 11 m | Pallas, Finland, PAL | 68.0° N, 24.1° E, 560 m |
| Bermuda, BME | 32.4° N, 64.4° W, 30 m | Ragged Pt, Barbados, RPB | 13.2° N, 59.4° W, 45 m |
| Cold Bay, Alaska, CBA | 55.2° N, 162.7° W, 25 m | Shemya Isl., Alaska, SHM | 52.7° N, 174.1° E, 40 m |
| Christmas Isl., CHR | 1.7° N, 157.2° W, 3 m | Ocean Stn, Norway, STM | 66.0° N, 2.0° E, 0 m |
| Mariana Isl., GMI | 13.4° N, 144.8° E, 1 m | Summit, Greenland, SUM | 72.6°, 38.5° W, 3238 m |
| Iceland, ICE | 63.3° N, 20.3° W, 118 m | Mt. Waliguan, China, WLG | 36.3° N, 100.9° E, 3810 m |
| Tenerife, IZO | 28.3° N, 16.5° W, 2360 m | Ny-Alesund, Sweden, ZEP | 78.9° N, 11.9° E, 475 m |
| Kumukahi, Hawaii, KUM | 19.5° N, 154.8° W, 3 m | | |

of a time-series can be found for datasets that contain gaps (Ghil et al., 2001). A new time-series, without gaps, can be reconstructed from these principle components. To generate the 'gap-filled' time-series, we fill the missing values in the original time-series with the estimates of these values from the reconstructed time-series. This process can be run iteratively, that is, the principle components of the 'filled' time-series are found and the points in the time-series that were missing in the original data are again filled with the new estimates. The 'gap-filled' time-series' at each station were then interpolated over latitude with a resolution of 5 degrees. This interpolated dataset has no longitudinal variation but is an observation based 'background' mixing ratio at 2-weekly resolution.

## 2.4  Bayesian inversion set-up

The N$_2$O mixing ratio at Ochsenkopf tall tower can be expressed simply in terms of the forward function of the state variables, that is, the atmospheric transport and mixing of the emissions. This function is provided by the surface influence term, $F(x_r, y_r, z_r, t_r | x_i, y_j, t_m)$ from STILT. Thus the forward model can be expressed in matrix form as:

$$c = \mathbf{F} \cdot \mathbf{f} + c_{\text{bnd}} + \varepsilon \qquad (2)$$

where $c$ is the measurement vector containing $m$ observations at the receptor point and $\mathbf{F}$ is a matrix ($m \times n$) in which each row is a vector containing the surface influence for each point $x_k$ (where $k = 1, \ldots, n$ and $n$ is the total number of state variables) over the time covered by the back-trajectory corresponding to the measurement with the same row index. In other words, each row of $\mathbf{F}$ relates the emissions, $\mathbf{f}$, to the mixing ratio in the corresponding row of the measurement vector, $c$. The second term on the right-hand-side, $c_{\text{bnd}}$, is a vector containing mixing ratios at the lateral boundary corresponding to the time of each observation and $\varepsilon$ is a vector containing the model-measurement mismatch. By rearrangement, Eq. (2) can be rewritten as:

$$y = \mathbf{F} \cdot \mathbf{f} + \varepsilon \qquad (3)$$

where $y$ is a vector with each element being the difference between the observed and the boundary mixing ratio, that is $y = c(x_r) - c_{\text{bnd}}$ and $\mathbf{F} \cdot \mathbf{f}$ is the forward model. The matrix $\mathbf{F}$ is defined according to the resolution of the retrieved fluxes, so that the impact of the flux in each pixel at each resolved time-step is related to the observation via each row of $\mathbf{F}$. The assumption is then made that the true fluxes, on which the observations depend, do not change significantly during the resolved time interval. This assumption leads to possible temporal aggregation errors, as the correction to the temporal variation in the emissions is aggregated into one emission field for each time-step and thus errors may result in their assignment (Law et al., 2004).

To test the impact of this assumption, we vary the temporal resolution for which the fluxes are retrieved. This corresponds to increasing the number of state variables in $\mathbf{f}$ by $1/l$, where $l$ is the temporal resolution of the resolved fluxes as a fraction of one year, and modifying $\mathbf{F}$ respectively so that it has the new dimensions of $m \times p$ (where $p = n/l$). However, there is a trade-off between spatial and temporal resolution; increasing the temporal resolution without decreasing the spatial resolution means an overall increase in the number of state variables, and hence a decrease in the possible constraint on each variable. Therefore, we chose to use a spatial resolution of $2 \times 2$ degrees, giving $n = 299$ for our domain.

Depending on the resolution, there may be more state variables than there are observations, that is, $p > m$, and the problem is under-determined. It can also be that in some time-steps a number of state variables (fluxes) will not have contributed at all to the mixing ratio at Ochsenkopf as a result of the atmospheric transport. In the Bayesian approach, an under-determined problem is still solvable because of the use of an a priori estimate, which is updated by the information

**Table 2.** Summary of the inversion tests and the parameters used.

|  | symbol | unit | A | B | C | D |
|---|---|---|---|---|---|---|
| true flux temporal variability |  | days | 7 | 1 | 1 | 1 |
| observation averaging interval |  | hours | 3 | 3 | 3 | 3–240 |
| time-step length |  | days | 7 | 7 | 7–30 | 7 |
| measurement error | $\sigma_{\text{meas}}$ | ppb | 0.001 | 0.1 | 0.1 | 0.1 |
| transport error | $\sigma_{\text{tran}}$ | ppb | 0 | 0.1 | 0.1 | 0.1 |
| boundary error | $\sigma_{\text{bnd}}$ | ppb | 0 | 0.1 | 0.1 | 0.1 |
| measurement error scale length | $M$ | days | 0 | 0.5 | 0.5 | 0.5 |
| transport error scale length | $A$ | days | 0 | 0.5 | 0.5 | 0.5 |
| boundary error scale length | $B$ | days | 0 | 30 | 30 | 30 |
| prior flux error for land fluxes | $\sigma_{\text{land}}$ | % | 100 | 100 | 100 | 100 |
| prior flux error for sea fluxes | $\sigma_{\text{sea}}$ | $\mu\,\text{mol}\,\text{m}^{-2}\,\text{s}^{-1}$ | 1e-06 | 1e-06 | 1e-06 | 1e-06 |
| prior flux spatial scale length | $D$ | km | 300 | 50–2000 | 300 | 300 |
| prior flux temporal scale length | $T$ | days | 30 | 7–30 | 30 | 30 |
| number of observations |  |  | 2880 | 2880 | 2880 | 2880–36 |
| number of state variables |  |  | 15 548 | 15 548 | 15 548–3588 | 15 548 |

contained in the observations (Enting, 2002; Rodgers, 2000; Tarantola, 2005). Obviously, where no information on a particular state variable is available, no constraint on the a priori estimate is possible. The uncertainty in the a priori estimate, however, must be quantified; this is achieved by specifying a probability density function over the state space. In this study, we assume that it is Gaussian. Similarly, the uncertainty in the observations, that is, the measurement error can be quantified as a probability density function over the measurement space. The best estimate of the state vector (i.e. that for which the cost function is minimized) can then be found from the Bayesian expression:

$$\mathbf{f} = \mathbf{f}_{\text{prior}} + \mathbf{S}_{\text{prior}} \cdot \mathbf{F}^{\text{T}} (\mathbf{F} \cdot \mathbf{S}_{\text{prior}} \cdot \mathbf{F}^{\text{T}} + \mathbf{S}_\varepsilon)^{-1} \cdot (\mathbf{y} - \mathbf{F} \cdot \mathbf{f}_{\text{prior}}) \quad (4)$$

where $\mathbf{S}_\varepsilon$ is the error covariance matrix of the measurements and $\mathbf{S}_{\text{prior}}$ is the error covariance matrix of the a priori estimate. This formulation of the inversion problem does not require finding the inverse of $\mathbf{S}_\varepsilon$ and $\mathbf{S}_{\text{prior}}$, thus making it numerically more efficient (the derivation and full description of Eq. (4) is provided by Tarantola et al., 2005). The pseudo-inverse of the matrix $(\mathbf{F} \cdot \mathbf{S}_{\text{prior}} \cdot \mathbf{F}^{\text{T}} + \mathbf{S}_\varepsilon)^{-1}$ was found using Singular Value Decomposition (SVD) (Press et al., 1992).

An accurate representation of the uncertainty in the a priori emissions is important as erroneous estimates, which have too narrow uncertainty bounds, can lead to estimated emissions that are inconsistent with the atmospheric observations and/or do not correspond to the true emissions patterns (Gerbig et al., 2006; Michalak et al., 2004). The uncertainty in the a priori emission estimates is represented in $\mathbf{S}_{\text{prior}}$, which has dimensions $p \times p$. The diagonal elements of $\mathbf{S}_{\text{prior}}$ are the squared errors for each of the $p$ state variables and the off-diagonal elements are the correlated errors between them. For the $p$ state variables, we used different errors for the variables corresponding to land and sea fluxes (see Table 2). The correlation was described by an exponential,

$\exp(-\Delta d/D - \Delta t/T)$ where $\Delta d$ is the distance between state variables and $\Delta t$ is the time interval between variables representing fluxes at the same location but at different points in time. The denominator $D$ is the spatial correlation scale length, and $T$ is the temporal correlation scale length. The choice of $D$ influences how strongly the different state variables are correlated to one another in space. For instance, a long correlation length, $D$ of the order of the domain dimensions, would mean a very strong correlation between state variables and would entail a strong dependence of the inversion on the spatial structure of the a priori emission estimate. On the other hand, a short $D$ of the order of one grid-cell's dimension would mean that the state variables are independent from one another, making the inversion less sensitive to the a priori spatial structure (Kaminski et al., 2001). The drawback of this, however, is that the error reduction on each state variable will be small and in the case where the observational constraint is weak there may be errors in the retrieval. As a starting point, we define $D$ to be 300 km, based on the scale of weather systems and land-cover, and $T$ to be 30 days so that errors in seasonal fluxes are uncorrelated. The influence of the choice of spatial correlation scale length has been previously investigated with regards to CO$_2$ (e.g. Gerbig et al., 2006). However, because the pattern of N$_2$O fluxes does not follow that of CO$_2$, we performed an analogous test to that of Gerbig et al. (2006) for N$_2$O and, additionally, tested the sensitivity of the inversion scheme to the choice of temporal correlation scale length (see Sect. 3.1.2).

The uncertainty in the observations is expressed in the measurement error covariance matrix, $\mathbf{S}_\varepsilon$, which is a square matrix of dimensions $m \times m$ with diagonal elements representing the error in each observation and off-diagonal elements representing the correlated errors between observations. The observations used in the inversion, $\mathbf{y}$, are the differences between the observed mixing ratios and the

contribution from advection of the lateral boundary mixing ratios. For this reason, the error in each observation, $y_i$, has contributions from the measurement, $\sigma_{\text{meas}}$, the estimate of the mixing ratio at the boundary, $\sigma_{\text{bnd}}$, and the transport model, $\sigma_{\text{trans}}$. The last error component, $\sigma_{\text{trans}}$, is included to account for imperfections in the transport model. The total error is thus:

$$\sigma_\varepsilon^2 = \sigma_{\text{meas}}^2 + \sigma_{trans}^2 + \sigma_{\text{bnd}}^2 \tag{5}$$

and $\sigma_\varepsilon^2$ is the value of the diagonal elements of $\mathbf{S}_\varepsilon$. The measurement, transport and boundary errors are assumed to be correlated over time. The transport error depends significantly on the accuracy of the Planetary Boundary Layer (PBL) height estimate, which varies throughout the day. The degree of correlation between transport errors is represented by an exponential function, $\exp(-\Delta t/A)$ where $\Delta t$ is the difference in time between measurements and $A$ is the temporal correlation scale length. We assumed that $A$ has a value of 0.5 day based on the temporal development of the PBL. The correlation of the boundary errors was treated in the same way using a temporal correlation scale length, $B$ of 30 days, which is approximately two times the temporal resolution of the boundary mixing ratio field. The temporal correlation of the measurement errors was treated likewise, using a correlation scale length of 0.5 day. Note that for the inversion tests, the value of each of the error components was the same as that assigned to the synthetic data (see Sect. 2.5)

## 2.5 Synthetic data generation

To test the inversion scheme, we generated a synthetic flux dataset of daily varying fluxes with a temporal correlation scale length of 30 days and a spatial correlation scale length of 300 km (see Fig. 1a and b). The synthetic fluxes were generated by adding a spatially and temporally varying offset to each grid-cell of a prior atemporal flux field (we used the GEIA N$_2$O flux estimates, Olivier et al., 1998). The offsets were calculated as the product of the square root of the error covariance matrix, $\mathbf{S}_{\text{prior}}^{1/2}$ (with 365$\times$n rows and columns) and a random number vector, $\mathbf{r}$ (of length 365$\times$n, with a zero mean and a SD of 1) after the method of Chevallier et al. (2007):

$$\mathbf{f}_{\text{syn}} = \mathbf{f}_{\text{prior}} + \mathbf{S}_{\text{prior}}^{1/2} \cdot \mathbf{r} \tag{6}$$

Because $\mathbf{S}_{\text{prior}}$ is too big to be stored at this temporal resolution, we avoided forming it directly and instead substituted $\mathbf{S}_{\text{prior}}^{1/2}$ in Eq. (6) using the following relationship:

$$\mathbf{S}_{\text{prior}}^{1/2} = \mathbf{S}_t^{1/2} \otimes (\mathbf{S}_s^{1/2} \cdot (\sigma^{1/2}(\sigma^{1/2})^{\text{T}})) \tag{7}$$

$\mathbf{S}_{\text{prior}}^{1/2}$ can be calculated as the Kronecker product (represented by $\otimes$) of $\mathbf{S}_t^{1/2}$ and $\mathbf{S}_s^{1/2} \cdot (\sigma^{1/2}(\sigma^{1/2})^{\text{T}})$, where $\mathbf{S}_t$ is the temporal correlation pattern and has dimensions 365$\times$365
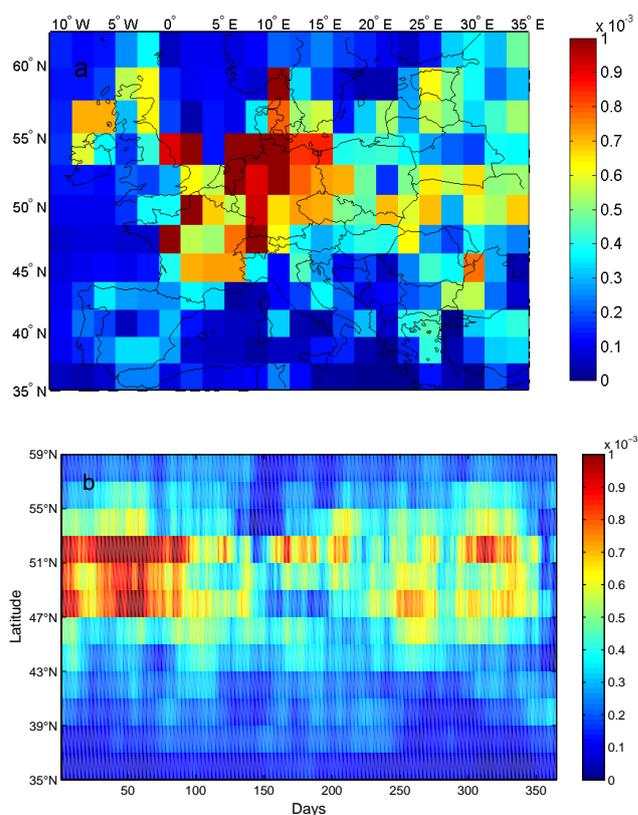


**Fig. 1.** The synthetic N$_2$O flux dataset that is used to generate the pseudo-observations. The synthetic fluxes contain spatially and temporally correlated errors according to the specifications of the prior error covariance matrix $\mathbf{S}_{\text{prior}}$: **(a)** the annual mean flux and **(b)** the zonal mean flux, both in units of $\mu$mol m$^{-2}$ s$^{-1}$

and $\mathbf{S}_s$ is the spatial correlation pattern and has dimensions $n \times n$ (both correlations patterns are described by an exponential decay function, see Sect. 2.3), and $\sigma$ is a vector containing the variances of the fluxes.

We coupled the synthetic fluxes with STILT-calculated surface influence functions to simulate atmospheric mixing ratios for the Ochsenkopf tall tower, which were used as the observations in the inversion tests. In tests B, C and D, we tested the influence that noise and data gaps in the observations have by adding Gaussian distributed noise of 0.1 ppb for each of the measurement, transport and boundary errors (quadratic sum = 0.2 ppb) with a temporal correlation scale length of 0.5 days for the measurement and transport errors, and 30 days for the boundary errors. Data gaps were introduced following the pattern of gaps in the real observations from the Ochsenkopf tall tower in which about 25% of the data are missing. Figure 2 shows the different time-series' where the 'true' mixing ratios (black line) are those generated from the synthetic fluxes (described above). The red line represents the mixing ratios that result from averaging the synthetic fluxes over 7-days, thus the difference between
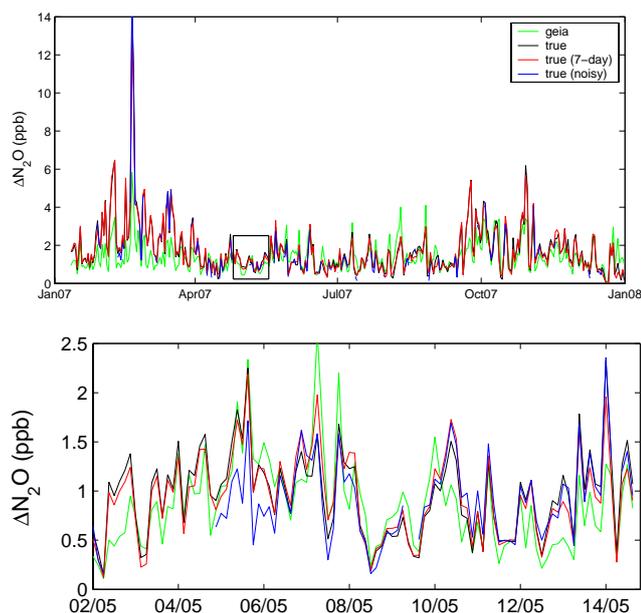
**Fig. 2.** Daily averaged mixing ratios generated using the STILT model coupled to the following flux maps: GEIA (green), true with daily variation (black), true averaged over 7-days (red), and true with added Gaussian distributed noise and data gaps (blue). A "zoom" of the section of the time-series' inside the frame is also shown at 3-hourly resolution.

the black and red lines is due to the influence of temporal aggregation in the fluxes, and the blue line shows the influence of adding Gaussian noise to the "true" mixing ratios. In all the inversion tests, we used the GEIA N$_2$O flux estimates (Olivier et al., 1998) as the a priori fluxes.

## 3 Results and discussion

### 3.1 Synthetic data tests

#### 3.1.1 Reference inversion

The validity of the inversion scheme was tested using the synthetic dataset described in Sect. 2.5. For this test, a temporal resolution of 7 days was used in the state space. The daily varying synthetic fluxes were averaged to this temporal resolution, to avoid aggregation errors, and mixing ratios were generated from the averaged fluxes using the STILT model. Since the same prior was used in the inversion, as was used to create the synthetic fluxes, the only source of error is the difference between the synthetic and the prior fluxes, which is known and is described by the prior error covariance matrix, $\mathbf{S}_{\text{prior}}$ (see Eq. 6). This test was used as a check for internal consistency and as a reference for following tests.

The performance of this inversion test, and all following tests, was assessed on the basis of the following criteria: (1)

the Root Mean Square Error (RMSE) of the retrieved and the synthetic (true) fluxes normalized by the RMSE of the prior and the true fluxes, (2) the reduced chi-squared value, which is equal to two times the cost-function at its minimum divided by the number of observations, and (3) the error reduction over the domain, calculated as $(1 - \sigma_{\text{post}}/\sigma_{\text{prior}})$ where $\sigma_{\text{post}}/\sigma_{\text{prior}}$ are the posterior/prior flux uncertainties taken as the square-root of the sum of covariances in the posterior/prior error covariance matrices according to Eq. (8), in which case $\sigma_{ij}$ is an element of $\mathbf{S}_{\text{post}}$:

$$\sigma_{\text{post}}^2 = \sum_{i=1, j=1}^{N} \sigma_{ij}^2 \qquad (8)$$

The posterior error covariance matrix, $\mathbf{S}_{\text{post}}$ was calculated according to Eq. (9) (Tarantola, 2005):

$$\mathbf{S}_{\text{post}} = \mathbf{S}_{\text{prior}} - \mathbf{S}_{\text{prior}} \cdot \mathbf{F}^{\text{T}} \cdot (\mathbf{F} \cdot \mathbf{S}_{\text{prior}} \cdot \mathbf{F}^{\text{T}} + \mathbf{S}_\varepsilon)^{-1} \cdot \mathbf{F} \cdot \mathbf{S}_{\text{prior}} \qquad (9)$$

In this reference test (A in Table 2), the RMSE of the fluxes was reduced by 48% and the reduced chi-square had a value of 0.81. The reduced chi-square can be thought of as the ratio of the actual error to the estimated error, therefore, when all errors are correctly accounted for in the inversion, it has an expected value of 1. In this reference test, the chi-square value is satisfactorily close to 1; this result, and the observed reduction in the RMSE of the fluxes, gives us confidence in the internal consistency of the inversion scheme.

#### 3.1.2 Sensitivity to temporal and spatial error correlations

All following tests in Sect. 3.1 were made using a similar set-up to the one used for the reference test but with a few important differences. The synthetic mixing ratios were generated by coupling STILT to the daily varying synthetic fluxes, while in the inversion the flux retrieval is at a temporal resolution of 7 days (unless otherwise stated), therefore they contain a component of temporal aggregation error. Furthermore, the mixing ratios were perturbed with Gaussian distributed noise (as described in Sect. 2.5) and data gaps were introduced at the same frequency as in the real observations. Details of each of the parameters used are given in Table 2.

The temporal and spatial correlation scale lengths used in the definition of the prior flux error covariance matrix have important implications for the retrieval (see Sect. 2.5). We tested the influence of changing the temporal correlation of errors in the state space by varying the correlation scale length, $T$, from 7 to 100 days (B in Table 2). The true value of $T$, which was used in the generation of the synthetic fluxes, is 30 days, therefore, for any given grid cell, the fluxes in consecutive time-steps will be correlated by $\exp(-7/30) = 0.79$. The minimum in the RMSE of the posterior and true fluxes was found for correlation lengths close to the true value of $T$ (see Fig. 3a). Increasing the correlation length, to example 100 days, means that fluxes in consecutive
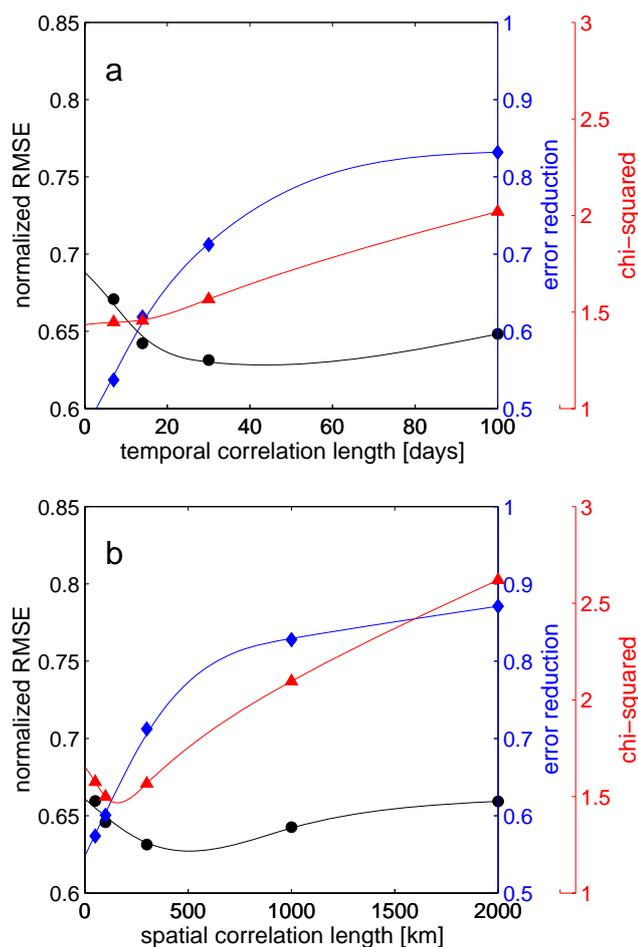
**Fig. 3.** The influence of changing the correlation scale lengths, used in the definition of the prior flux error covariance matrix, on the Root Mean Square Error (RMSE) of the posterior and true fluxes (black circles), the error reduction (blue diamonds) and the reduced chi-square value (red triangles), for **(a)** changing the temporal correlation scale length and **(b)** changing the spatial correlation scale length. All these figures show the results of inversions using synthetic mixing ratios with data gaps and added Gaussian noise. The lines are cubic spline fits to the data.

time-steps would be correlated by 0.93, which is a significant over-estimate of the actual correlation, and hence there is reduction in the freedom to correct the prior temporal variability towards that of the true temporal variability. On the other hand, if the correlation length is under-estimated, then there is too little constraint on the fluxes in different time-steps. Similarly, the error reduction increases with increasing $T$, since the information used to constrain one time-step has an influence on other time-steps over longer intervals. This means that in the case where $T$ is over-estimated, there is also a corresponding under-estimation of the posteriori error and hence an over-estimation of the error reduction. The reduced chi-square value also increases with increasing correlation length, since the actual error (that is between the observed

and modeled mixing ratios and between the prior and posterior fluxes) increases relative to the prescribed error in the inversion set-up.

In an analogous test, the influence of changing the spatial correlation length, $D$, from 50 to 2000 km was examined (see Fig. 3b). Again, the RMSE is at a minimum for values of $D$ close to the true value (300 km). The reduced chi-square increases for spatial correlation lengths longer than the true length, similar to the result for the test on temporal correlation, as the error between the observed and modeled mixing ratios increases. However, using correlation lengths shorter than the true value of $D$ also leads to an increase in the reduced chi-square, as this results in a smaller prior error (in $\mathbf{S}_{\text{prior}}$) and thus a larger ratio of the actual to prescribed errors.

The tests presented indicate the sensitivity of the inversion to the choice of correlation scale length; however, they are limited in the sense that they do not help to determine the true correlation scale length. This is an area of ongoing research and goes beyond the scope of this paper.

The value of the reduced chi-square for the case when $T$ = 30 days and $D$ = 300 km is noteworthy as it has increased from 0.81 in the reference inversion case (Sect. 3.1.1) to 1.56, even though the Gaussian errors in the mixing ratios have been accounted for. This increase is due to the component of temporal aggregation error, which arises from inverting mixing ratios that contain a signal from the daily flux variations, at a 7-day temporal resolution.

### 3.1.3 Temporal resolution of the resolved fluxes and aggregation errors

The impact of temporal aggregation error was tested by performing inversions with varying flux resolution (C in Table 2). Temporal aggregation errors are incurred when the fluxes are changing significantly from the mean flux over the resolved time interval. Since the observations used in our inversion tests were generated from daily varying fluxes, whereas the retrieved fluxes were resolved only at 7-day and longer intervals (the resolutions tested were 7, 10, 14 and 30 days), there is a component of aggregation error in all these inversions. The influence of the aggregation error can be seen in the normalized RMSE and reduced chi-square value, which increase with the length of the resolved time interval (see Fig. 4a). This behaviour is similar to that when changing the temporal correlation scale length; one assumes that the fluxes within the resolved time interval have a correlation of 1, compared with the true correlation e.g. at 30 days which is 0.37. However, the correlation between the time intervals decays according to the prescribed function, in our case: $\exp(-\Delta t/T)$. The aggregation error can be projected into the observation space, where it can be thought of as the difference between the true mixing ratios, which would be obtained from the daily varying fluxes, and the mixing ratios that can be obtained from the lower resolution fluxes (e.g.
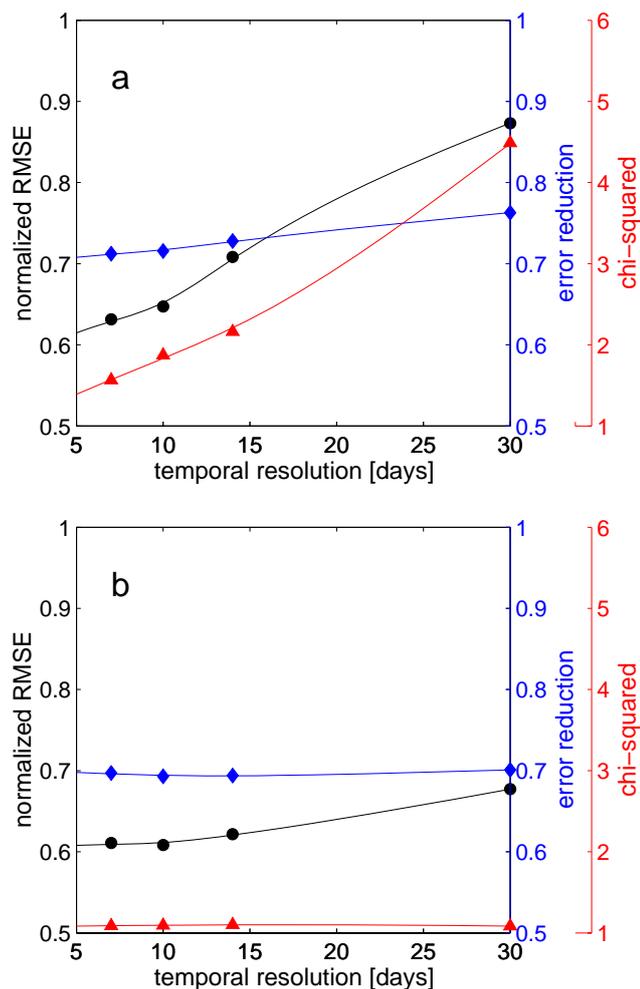
**Fig. 4.** Analysis of the impact that temporal resolution in the state space has on the retrieval. Shown is the influence on RMSE of the posterior and true fluxes (black circles), error reduction (blue diamonds) and the reduced chi-square value (red triangles) for two cases **(a)** using errors in the observation space with contributions from only transport, measurement and boundary errors and **(b)** including in addition the estimated aggregation error for each temporal resolution. The lines are cubic spline fits to the data.

see Fig. 2). Moreover, since increasing the averaging interval results in an additional error, which is not accounted for in the observation error covariance matrix, there is an under-estimation of the prior error and thus a corresponding increase in the reduced chi-squared value.

Kaminski et al. (2001) address the problem of spatial aggregation error and proposed an algorithm (based on that of Trampert et Snieder, 1996) for estimating this error. Their algorithm is based on propagating the information lost by reducing the spatial resolution of the retrieved fluxes (that is with respect to that of the transport model used in the inversion) and projecting this into the measurement space. Although Kaminski et al. (2001) do not discuss temporal aggre-

gation errors, their algorithm can also be applied in the same way to propagate the error induced by reducing the temporal resolution of the inversion. The algorithm involves two steps: (1) calculation of the prior flux uncertainty that is not resolved in the lower resolution model and (2) projection of this uncertainty into the observation space (see the appendix for a detailed description).

We calculated the additional error in the observation space incurred by reducing the temporal resolution of the resolved fluxes from 1 day, coinciding with the variability in the true fluxes, to that used in the inversion. Inversions were performed using the same range of temporal resolutions as in the above test (see Fig. 4b). In contrast to the test that did not account for the aggregation error, the normalized RMSE only increased by approximately 0.1 over the range of temporal resolutions. Furthermore, the reduced chi-square value remained nearly constant over the whole range of resolutions, with all values between 1.08 and 1.10, reflecting the fact that the estimated errors, which included the aggregation error, were close to the actual errors. The magnitude of the aggregation error calculated for the observation space (mean over all observations) was 0.22 ppb and 0.49 ppb for resolutions of 7 and 30 days, respectively, compared with a total observational error (i.e. the quadratic sum excluding the aggregation error) of 0.2 ppb, and thus represents a significant contribution to the overall observational error.

### 3.1.4 Averaging interval of the observations

Another proposed method for dealing with temporal aggregation error, is to average the observations over the same time interval as the retrieved fluxes, and thereby smooth out signals which cannot be represented at the temporal resolution of the model and thus reduce the bias errors (Peylin, 2002). We tested the influence of increasing the averaging interval of the observations up to 240 h (D in Table 2).

From Fig. 5 it can be seen that averaging in the observation space has a rather complex affect on the retrieval. The best performing inversion, in terms of the RMSE, is the one that uses observations at the highest resolution. However, there is also a small secondary RMSE minimum at 168 hours, which coincides with the averaging interval of the fluxes (7 days). This result can be understood in that by averaging the observations, the surface influence is also averaged and is hence spread-out over a larger region, therefore, averaging the observations means that the ability of the inversion to resolve the fluxes in different grid-cells is reduced. This result holds in the case where there are no transport errors. The secondary RMSE minimum at 168 hours, corresponding to the flux resolution in the inversion, results from the fact that averaging the observations reduces the influence of aggregation error in the fluxes. The chi-square value is determined by 3 factors: the posterior-prior flux difference ($\mathbf{f}_{post}$-$\mathbf{f}_{prior}$), the error in the observation space ($\mathbf{Ff}_{post} - \mathbf{y}$), and the magnitude of the prescribed errors for both of these. The posterior-prior
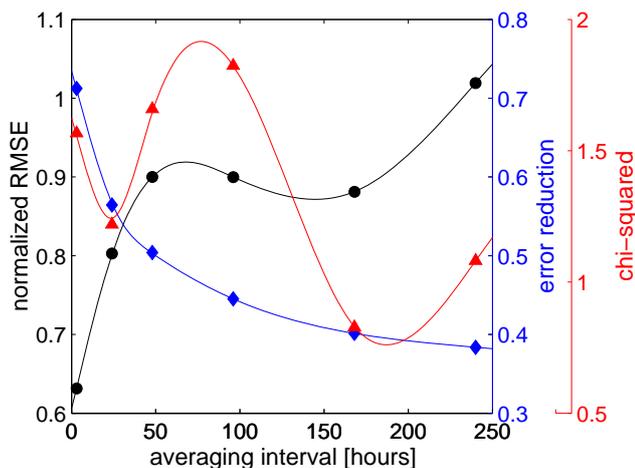
**Fig. 5.** The influence that changing temporal resolution in the observation space has on the retrieval (using a fixed resolution in the state space of 7-days). Shown is the impact on the RMSE of the posterior and true fluxes (black circles), error reduction (blue diamonds) and the reduced chi-square value (red triangles). The lines are cubic-spline fits to the data.

flux difference is largest for the 3-hourly observations and decreases with averaging interval length, whereas the opposite is true for the observation errors. However, the shape of the chi-square curve is not quite the balance of these two, as it also depends on the error estimates. The minimum at 24 h is likely due to the prior assumption of the temporal correlation scale length in the observation error covariance matrix, while the minima at 168 hours (close to 1) is likely because the prior error estimate is close to the actual error, since there is no aggregation error at this time. This shows that the chi-square, although a useful parameter for assessing the inversion, is by itself not always a sufficient one. The error reduction also decreases with increasing averaging interval, as there are fewer independent observations in the inversion to constrain the fluxes.

## 3.2 Real observation inversion

### 3.2.1 N₂O flux estimates

N₂O fluxes were retrieved for 2007 using the atmospheric mixing ratios observed at the Ochsenkopf tall tower. The inversion set-up was the same as that used in the reference inversion with synthetic data with a few exceptions. For the a priori fluxes we used the IER dataset (http://carboeurope.ier.uni-stuttgart.de) for the anthropogenic fluxes and GEIA for fluxes from the ocean and from soils under natural vegetation, the combination of which we believe is a closer estimate to the true fluxes. However, we kept the same temporal and spatial correlation scale lengths, that is, 30 days and 300 km, respectively, and the same temporal resolution of 7 days, as were used in the reference inversion case. The mixing ratios

at the boundaries were taken from the fields of interpolated NOAA flask data as described in Sect. 2.3. In the observation space, we used measurements at 3-hourly intervals, each with an uncertainty of 0.5 ppb, which is the quadratic sum of squares of the measurement (0.3 ppb), transport (0.3 ppb), boundary (0.2 ppb), and aggregation errors (0.2 ppb). The measurement error was estimated on the basis of the long-term precision (see Sect. 2.1) and comparisons of the in-situ with independent flask measurements (Thompson et al. 2009). For the transport error, we used a simple approximation based on the mean of the 3-day SD's of the simulated N₂O mixing ratio at Ochsenkopf. The reason for this is that transport models are generally not able to represent synoptic scale variability well, which typically has time spans of circa 3-days. The boundary error was estimated from the SD the boundary contribution to the mixing ratio observed at Ochsenkopf. Although the boundary mixing ratios likely under-estimate the real variability (because they are based on the interpolation of bi-weekly observations at 5 degrees meridional resolution and no longitudinal resolution) the impact on the simulated mixing ratios at Ochsenkopf is likely to be small, since the total variability at this site is dominated by processes occurring within the domain. Using this set-up, we achieved a posterior reduced chi-square value of 1.17, which to some degree indicates an appropriate choice of a priori and measurement uncertainties and covariances.

To verify the posterior fluxes, we coupled them to the STILT model to simulate mixing ratios at the Ochsenkopf site as well as for an independent flask sampling site, Biaystok (53°13′ N, 23°1′ E), which is in the MPI-BGC network (see Fig. 6). These simulated mixing ratios were then compared with observations at each site. A significant reduction in the RMSE of the posterior and observed mixing ratios was seen at Ochsenkopf, as well as improvements in the correlation coefficient, normalized standard deviation and regression coefficient (see Table 3). Similarly, at Bialystok, a small improvement in the fit to observations was seen using the posterior fluxes, although there was still an under-estimate of the observed mixing ratios in August and September, owing to the weaker constraint on fluxes in the vicinity of Bialystok from the Ochsenkopf site. In Fig. 6 the contribution to N₂O variability from transport at the domain boundaries is also shown (green line). At each of these sites, the contribution from the boundary defines the trend and seasonal cycle of N₂O, whereas the synoptic variability is predominantly due to effects inside the domain. Since the boundary mixing ratios are based on observations, errors in boundary mixing ratios are mostly due to transport errors in the STILT model and from the interpolation of the sparse data.

### 3.2.2 Posterior error and error reductions

The posterior variances for each pixel are the diagonal elements of the error covariance matrix $\mathbf{S}_{post}$ calculated
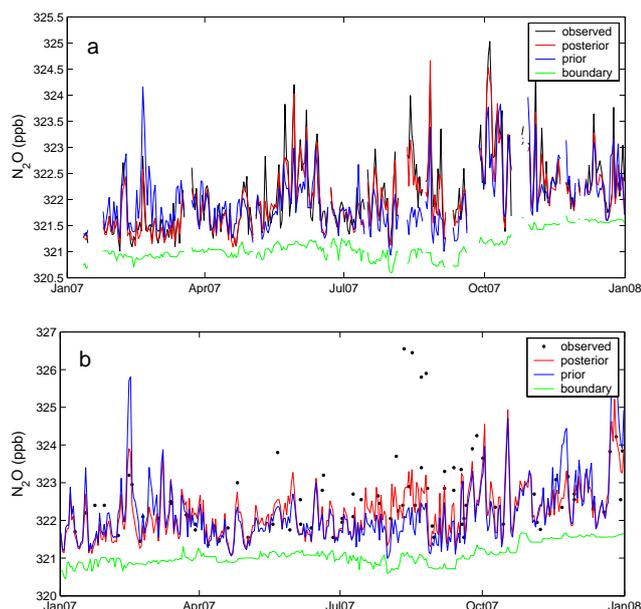
**Fig. 6.** Comparison of the observed (black), posterior (red), prior (blue) and boundary mixing ratios (green). The boundary mixing ratio is the mixing ratio at the receptor that would be seen in the absence of any fluxes inside the domain, in other words, it is the N$_2$O that is transported from the domain boundaries to the measurement site. Mixing ratios are shown for **(a)** the Ochsenkopf tall tower, and **(b)** for the MPI flask site Bialystok (53°13′ N, 23°1′ E).

**Table 3.** Comparison of prior and posterior mixing ratios at Ochsenkopf and Bialystok.

|  |  | RMSE (ppb) | $R^2$ | normalized SD | regression coefficient |
|---|---|---|---|---|---|
| Ochsenkopf | prior | 0.53 | 0.44 | 0.83 | 0.75 |
|  | posterior | 0.28 | 0.85 | 0.91 | 0.91 |
| Bialystok | prior | 1.13 | 0.16 | 0.79 | 0.57 |
|  | posterior | 0.88 | 0.37 | 0.66 | 0.52 |

according to Eq. (9) and are dependent on the assumptions made for the prior error covariance matrix **S**$_{prior}$ (i.e. the correlation scale length and variances), the observation error covariance matrix S$_\varepsilon$, as well as on the surface influence function. Therefore, the posterior error (and error reduction) reflects not only the constraint provided by the atmospheric observations but also the parameterization of the inversion framework, thus these results should not be over interpreted. The spatial extent of the error reduction depends strongly on the spatial correlation length used in the inversion. We chose to use a correlation length of 300 km; although higher values would have resulted in a greater spread of error reduction, an overestimate of this parameter could lead to wrong assumptions about the correlation of flux errors, and thus, errors in the retrieval.

Although only one in-situ measurement site was used in the inversion, substantial error reductions were seen in western and central Europe, between 30 and 60%, with the highest reductions in Germany close to the Ochsenkopf site (see Fig. 7d). The pattern of error reduction closely follows that of the annual mean surface influence function (or footprint) as calculated for Ochsenkopf using STILT (footprint not shown). This result is as expected, since the regions that have the strongest influence on the observed mixing ratios, owing to atmospheric transport, will be the best constrained in the inversion, whereas regions that have only a weak influence on the observations are poorly constrained.

### 3.2.3 Temporal variability of the fluxes

The observational constraint at Ochsenkopf resulted in a notable reduction in N$_2$O fluxes, relative to the prior, in February and March for Germany and Poland, which for Germany extended into April. In May, however, this trend reversed and there was a significant increase in fluxes over most of Europe, which was sustained until October. The increase was particularly large in August and September (see Fig. 8).

To better understand what was driving the change in N$_2$O flux, we looked at the temporal variability of the mean fluxes over 6 regions in Europe (northern France, southern France, northern Germany, southern Germany, United Kingdom, and Poland) and compared these with regional meteorological parameters (see Fig. 8). We performed a linear multiple analysis of variance of the 7-day posterior fluxes with the 7-day normalized anomalies in temperature, soil moisture and rainfall, where these were calculated using the mean values for each region based on ECMWF analysis fields. Although this approach looks at the integrated influence of these parameters over large regions, some significant correlations were seen; the variance with temperature anomaly was significant for all regions at the 95% confidence level, whereas rainfall was only significant for southern Germany and northern France, and soil moisture was only found to be significant in the UK. We found a positive correlation with temperature anomaly at all sites varying between 0.4 and 0.8 (see Table 4). In southern Germany and northern France, N$_2$O fluxes were also positively correlated with rainfall. In the UK, N$_2$O fluxes were negatively correlated with soil moisture, a result that can be understood when the soil water filled pore space (WFPS) is considered. N$_2$O fluxes have been shown to increase with increasing soil moisture up to approximately 80% WFPS, however, above this value N$_2$O emissions decrease sharply, thus in already very moist soils, increasing soil moisture leads to a decrease in N$_2$O flux (Freibauer and Kaltschmitt, 2003). The differences between regions for the result of correlation with rainfall and soil moisture depend strongly on other factors not considered in this analysis, such as soil type, oxygen and nitrate substrate availability, as well as on land management (e.g. Freibauer and Kaltschmitt, 2003). Although it is outside the scope of this
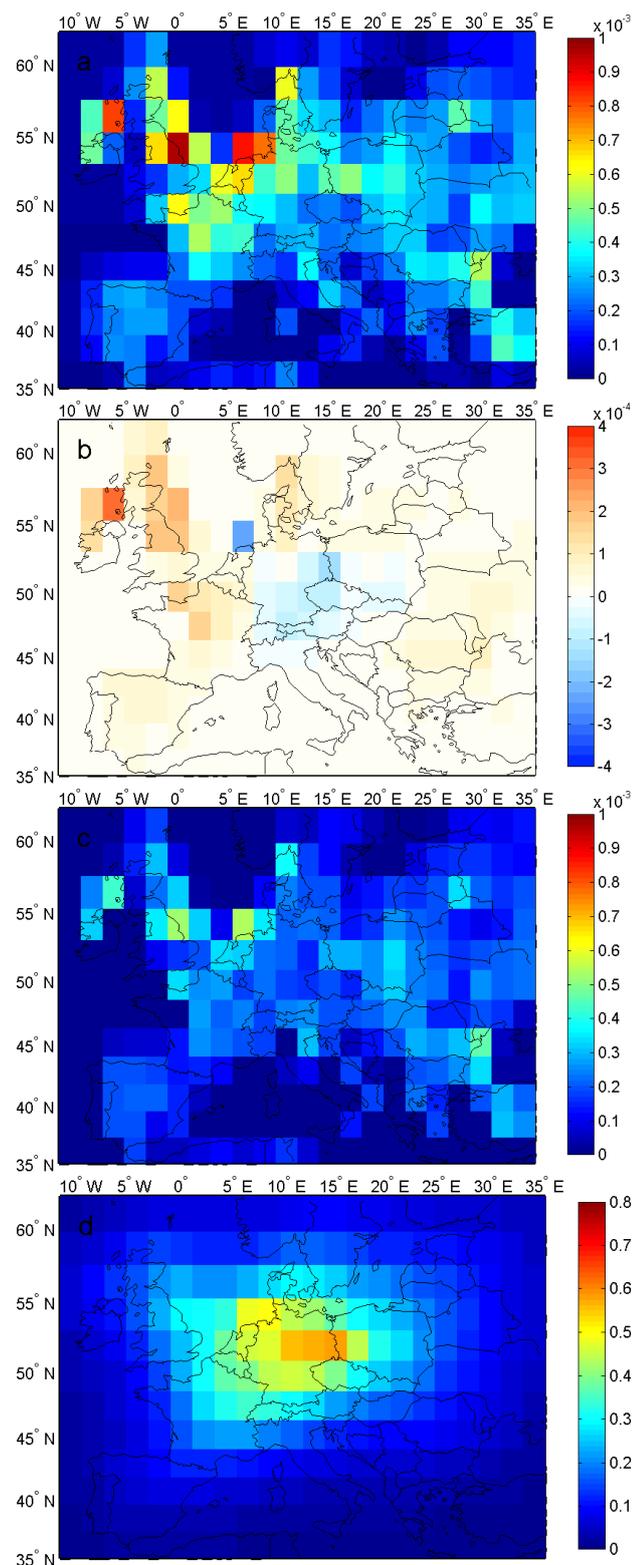
**Fig. 7.** Inversion results using the observed N$_2$O mixing ratio at Ochsenkopf: **(a)** annual mean posterior N$_2$O flux, **(b)** annual mean difference between the posterior and prior N$_2$O flux, c) posterior error for the annual mean flux, and d) error reduction calculated as $1-\sigma_{post}/\sigma_{prior}$. All fluxes are in units of $\mu$mol m$^{-2}$ s$^{-1}$.
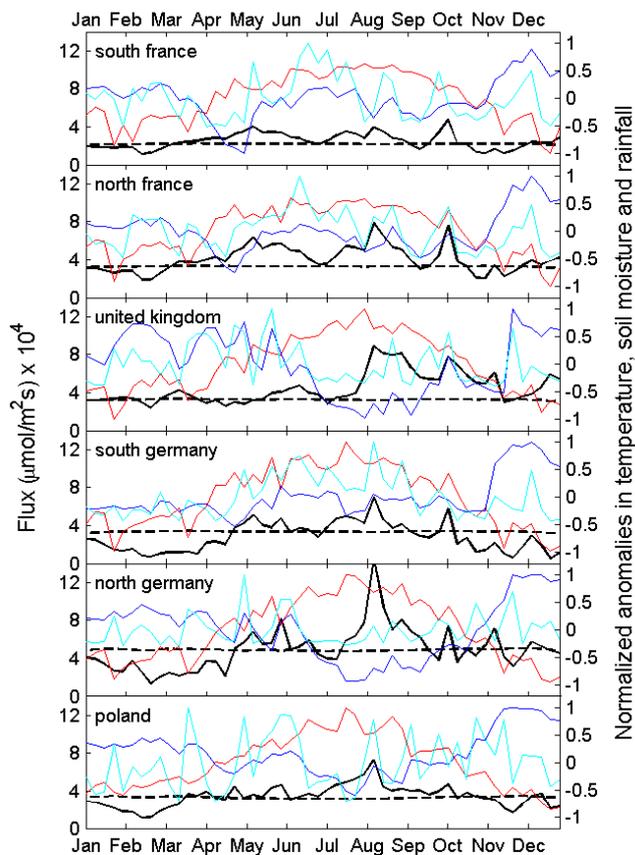


**Fig. 8.** Temporal variability of the total posterior (black solid) and prior (black dashed) N$_2$O flux for 6 regions in Europe compared with the normalized anomalies in mean temperature (red), soil moisture (blue) and rainfall (cyan) for each region in 2007.

**Table 4.** Correlations of N$_2$O fluxes with temperature, soil moisture and rainfall calculated by multiple analysis of variance,

|  | temperature | soil moisture | rainfall |
|---|---|---|---|
| southern France | 0.54 | N.S. | N.S. |
| northern France | 0.63 | N.S. | 0.42 |
| UK | 0.43 | −0.51 | N.S. |
| southern Germany | 0.80 | N.S. | 0.70 |
| northern Germany | 0.57 | N.S. | N.S. |
| Poland | 0.71 | N.S. | N.S. |

N.S. = Not Significant (at the 95% confidence level).

study to examine these, we recognise that activities such as the application of N-fertilizer and soil tillage also have a significant impact on N$_2$O fluxes (Li et al., 1996). N-fertilized applications may contribute to the high emissions seen between April and June, while tillage and harvest may contribute to the high emissions between August and September.

**Table 5.** Comparison of annual mean N$_2$O emissions for five European countries in units of Gg-N$_2$O/y.

|  | posterior[1] | prior[2] | EDGAR[3] | Manning et al.[4] |
|---|---|---|---|---|
| France | 260±144 | 214 | 139 | 271, (252–287) |
| Germany | 225±135 | 242 | 238 | 247, (220–274) |
| UK | 150±87 | 111 | 79 | 121, (107–134) |
| Benelux | 72±44 | 78 | 98 | 101, (81–146) |
| Poland | 151±116 | 144 | 77 | not given |

1. the total uncertainties were calculated by integrating the posterior uncertainty in each grid-cell over the land surface area (where the grid-cell uncertainty was calculated according to Eq. 8).
2. the prior used in this study included the anthropogenic emissions of IER (http://carboeurope.ier.uni-stuttgart.de) and the natural soil and ocean emissions of GEIA, Olivier et al. (1998) with uncertainties of 100%.
3. EDGAR-32FT2000 (anthropogenic emissions only) with a nominal uncertainty of 100%, Olivier et al. (2005).
4. uncertainty range given inside parentheses, Manning et al. (2003).

### 3.2.4 Annual total fluxes and their comparison with other estimates

The main information about the fluxes, provided by the atmospheric observations in the inversion, is on their temporal variability as discussed in Sect. 3.2.3. Conversely, the annual mean posterior flux estimate is close to that of the prior (see Fig. 7b). The largest increments made to the fluxes (i.e. the largest annual mean posterior-prior differences) were in fact in pixels that were not well constrained by the observations and, thus, where the error reductions were low, only between 10 and 20% (see Fig. 7d). These results, therefore, should not be over-interpreted, specifically, the results for the northern UK where very large fluxes have been retrieved in the inversion. Reasons for this could be due to the limited resolving power of the model; in order to reduce the model – observation mismatch, increased fluxes are required from the region to the northwest of Ochsenkopf, however, the distribution of this additional flux is not well constrained and depends not only on the surface influence function but also on the prior flux uncertainties.

The annual mean N$_2$O emission, and its uncertainty, of 5 countries/regions within our domain (Benelux, France, Germany, Poland, and UK) were calculated and compared to the estimates of IER and GEIA (i.e. our prior estimate), EDGAR-v3.2, and Manning et al. (2003). The annual totals a posteriori, a priori, and as estimated in EDGAR-v3.3 and by Manning et al., all agreed within the given uncertainties. However, the flux estimates of EDGAR-v3.2 for France and Poland had the largest differences from other mean estimates but were still within the uncertainty range (see Table 5).

## 4 Summary and conclusions

In this study, we have firstly examined how high temporal resolution atmospheric measurements of N$_2$O may be best incorporated into an inversion scheme, while regarding the potential errors that may arise, and secondly retrieved N$_2$O emissions estimates for 2007 for central and western Europe. By using synthetic datasets with known errors both in the observations and in the fluxes, we explored the influence of temporal and spatial correlation length as well as the temporal resolution of the fluxes. From these tests we find that temporal aggregation of the fluxes can lead to significant errors in the retrieval if not accounted for. We found for the pseudo-data case, where the true fluxes are varying daily (to which the observations are sensitive), using temporal resolutions of 7-days and longer in the retrieval leads to significant aggregation errors. These errors can be accounted for by increasing the observational error, so that in addition to errors in the transport model, measurements and boundary conditions, the errors resulting from the loss of information by going to a lower resolution model (the aggregation errors) are included. The algorithm of Kaminski et al. (2001) can be used to estimate the aggregation error in the observation space and only requires an estimate on what timescale the true fluxes are varying. Although this study only included one observation site, the algorithm can be employed to calculate aggregation errors at multiple sites and the investigation into the importance of aggregation errors for such an observation network constitutes an important next step. Another result of this work that needs to be investigated for a network is the influence of averaging the observations in time and whether using higher frequency data would improve the retrieval of the fluxes, as was found for the synthetic test D in this study.

Using the observations from the tall tower site, Ochsenkopf (50°01′ N, 11°48′ E) we retrieved N$_2$O fluxes at a 7-day resolution for 2007. The posterior fluxes showed much greater temporal variability than the prior fluxes. Relative to the prior, notable reductions in N$_2$O fluxes were seen in February and March for Germany and Poland, while in May, a significant increase was seen over most of Europe, which was sustained until October. In particular, large emissions were seen in August and September, which may be due to human activities such as harvest and tillage at this time. Over the whole year and in all regions, we found N$_2$O fluxes to be significantly correlated (at the 95% confidence level) with temperature anomaly, while correlations with rainfall were only significant for southern Germany and northern France, and correlations with soil moisture were only significant for the UK. The observational constraint had the most impact on the temporal variability of the fluxes, while the annual mean fluxes remained close to those of the prior (see Table 5). Moreover, the mean fluxes found for France, Germany, the UK and Benelux, agreed very well with those found in the inversion of Manning et al. (2003)

and the estimates from EDGAR-32FT2000 also fell within the posterior uncertainty range. Despite using only one observation site, substantial error reductions of 30–60% were achieved for fluxes in central and Europe.

## Appendix A

## Description of the algorithm for estimating aggregation errors

We present a description of the algorithm derived by Kaminski et al. (2001) and our method for its calculation. The first step involves defining a projection operator, $\mathbf{P}_+$ that operates in the space of the high-resolution model and projects only the components that are resolved by the low-resolution model:

$$\mathbf{P}_+ = \Sigma(\boldsymbol{p}_i \boldsymbol{p}_i^{\mathrm{T}}) \tag{A1}$$

where $\boldsymbol{p}_i$ are vectors expressing the mapping of the high to the low resolution model and have unit length and are orthogonal. The remaining unresolved components are thus defined by:

$$\mathbf{P}_- = \mathbf{I} - \mathbf{P}_+ \tag{A2}$$

where $\mathbf{I}$ is the identity matrix. In the second step, the aggregation error covariance matrix, $\mathbf{S}_{\mathrm{agg}}$ is calculated by projecting the error in the flux space into the observation space:

$$\mathbf{S}_{\mathrm{agg}} = \mathbf{F} \cdot \mathbf{P}_- \cdot \mathbf{S}_{\mathrm{prior}} \cdot \mathbf{P}_-^{\mathrm{T}} \cdot \mathbf{F}^{\mathrm{T}} \tag{A3}$$

where the footprint, $\mathbf{F}$ and prior flux error covariance matrix, $\mathbf{S}_{\mathrm{prior}}$ are defined for the high resolution model. Since $\mathbf{P}_-$ is too large to be calculated directly (in our case $\mathbf{P}_-$ would be a $100\,000 \times 100\,000$ matrix) we calculate directly the product, $\mathbf{F} \cdot \mathbf{P}_-$ (which is approximately $2500 \times 100\,000$):

$$\mathbf{F} \cdot \mathbf{P}_- = \mathbf{F} - \mathbf{F} \cdot \Sigma(\boldsymbol{p}_i \boldsymbol{p}_i^{\mathrm{T}}) \tag{A4}$$

Here, our method departs from that of Kaminski et al. (2001); we calculate the product $\mathbf{F} \cdot \mathbf{P}_- \cdot \mathbf{S}_{\mathrm{prior}}$ and thus avoid forming $\mathbf{S}_{\mathrm{prior}}$ directly (using the same method described in Sect. 2.4):

$$\mathbf{F} \cdot \mathbf{P}_- \cdot \mathbf{S}_{\mathrm{prior}} = \mathbf{F} \cdot \mathbf{P}_- \cdot \mathbf{S}_t \otimes (\mathbf{S}_s \cdot (\sigma(\sigma)^{\mathrm{T}})) \tag{A5}$$

and lastly $\mathbf{S}_{\mathrm{agg}}$ is calculated according to Eq. (A5), which is equivalent to Eq. (A3):

$$\mathbf{S}_{\mathrm{agg}} = \mathbf{F} \cdot \mathbf{P}_- \cdot \mathbf{S}_{\mathrm{prior}} \cdot (\mathbf{F} \cdot \mathbf{P}_-)^{\mathrm{T}} \tag{A6}$$

The publication of this article is financed by CNRS-INSU.

## References

Biraud, S., Ciais, P., Ramonet, M., Simmonds, P., Kazan, V., Monfray, P., O'Doherty, S., Spain, G., and Jennings, S. G.: Quantification of carbon dioxide, methane, nitrous oxide and chloroform emissions over Ireland from atmospheric observations at Mace Head, Tellus, 54B, 41–60, 2002.

Bremner, J. M.: Sources of nitrous oxide in soils, Nutr. Cycl. Agroecosyst., 49(1–3), 7–16, 1997.

Chevallier, F., Bréon, F.-M. and Rayner, P. J.: Contribution of the Orbiting Carbon Observatory to the estimation of CO2 sources and sinks: Theoretical study in a variational data assimilation framework, J. Geophys. Res., 112, D09307, doi:10.1029/2006JD007375, 2007

Corazza, M., Bergamaschi, P., Vermeulen, A. T., Aalto, T., Haszpra, L., Meinhardt, F., O'Doherty, S., Thompson, R., Moncrieff, J., Popa, E., Steinbacher, M., Jordan, A., Dlugokencky, E., Brhl, C., Krol, M., and Dentener, F.: Inverse modelling of European N2O emissions: assimilating observations from different networks, Atmos. Chem. Phys., 11, 2381–2398, doi:10.5194/acp-11-2381-2011, 2011.

Crutzen, P. J.: Estimates of Possible Variations in Total Ozone Due to Natural Causes and Human Activities, Ambio, 3(6), 201–210, 1974.

Enting, I. G.: Inverse Problems in Atmospheric Constituent Transport, Cambridge University Press, Cambridge, New York, USA, 408 pp., 2002.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D. W., Haywood, J., Lean, J., Lowe, D. C., Myhre, G., Nganga, J., Prinn, R., Raga, G., Schultz, M. and Van Dorland, R.: Changes in Atmospheric Constituents and in Radiative Forcing, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L., Cambridge University Press, Cambridge, UK, 129–234, 2007.

Freibauer, A. and Kaltschmitt, M.: Controls and models for estimating direct nitrous oxide emissions from temperate and subboreal agricultural mineral soils in Europe, Biogeochemistry, 63, 93–115, 2003.

Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P. and Sut-

ton, M. A.: Transformation of the Nitrogen Cycle: recent trends, questions, and potential solutions, Science, 320, 889–892, 2008.

Gerbig, C., Lin, J. C., Munger, J. W. and Wofsy, S. C.: What can tracer observations in the continental boundary layer tell us about surface-atmosphere fluxes?, Atmos. Chem. Phys., 6, 539–554, doi:10.5194/acp-6-539-2006, 2006.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bawkin, P., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO$_2$ with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-orientated framework, J. Geophys. Res., 108(D24), 4757, doi:10.1029/2003JD003770, 2003a.

Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Andrews, A. E., Stephens, B. B., Bawkin, P. S., and Grainger, C. A.: Toward constraining regional-scale fluxes of CO$_2$ with atmospheric observations over a continent: 1. Observed spatial variability from airborne platforms, J. Geophys. Res., 108(D24), 4756, doi:10.1029/2002JD003018, 2003b.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A. W., Saunders, A., Tian, Y., Varadi, F. and Yiou, P.: Advanced Spectral Methods for Climatic Time Series, Rev. Geophys., 40(1), 1-1–1-41, doi:10.1029/2001RG000092, 2001.

Gruber, N. and Galloway, J. N.: An Earth-system perspective of the global nitrogen cycle. Nature, 451, 293–296, doi:10.1038/nature06592, 2008

Hirsch, A. I., Michalak, A. M., Bruhwiler, L. M., Peters, W., Dlugokencky, E. J. and Tans, P. P.: Inverse modeling estimates of the global nitrous oxide surface flux from 1998–2001, Global Biogeochem. Cy., 20, GB1008, doi:10.1029/2004GB002443, 2006.

Huang, J., Golombek, A., Prinn, R., Weiss, R., Fraser, P., Simmonds, P., Dlugokencky, E. J., Hall, B., Elkins, J., Steele, P., Langenfelds, R., Krummel, P., Dutton, P., and Porter, L.: Estimation of regional emissions of nitrous oxide from 1997 to 2005 using multinetwork measurements, a chemical transport model, and an inverse method, J. Geophys. Res., 113, D17313, doi:10.1029/2007JD009381, 2008

Kaminski, T., Rayner, P. J., Heimann, M., and Enting, I. G.: On aggregation errors in atmospheric transport inversions, J. Geophys. Res., 106(D5), 4703–4715, 2001.

Kort, E. A., Eluszkiewicz, J., Stephens, B. B., Miller, J. B., Gerbig, C., Nehrkorn, B. C., Daube, B. C., Kaplan, J. O., Houweling, S., and Wofsy, S. C.: Emissions of CH$_4$ and N$_2$O over the United States and Canada based on a receptor oriented modeling framework and COBRA-NA atmospheric observations, Geophys. Res. Lett., 35, L18808, doi:10.1029/2008GL034031, 2008

Kroeze, C., Dumont, E., and Seitzinger, S. P.: New estimates of global emissions of N$_2$O from rivers and estuaries, Environ. Sci., 2, 159–165, 2005.

Law, R. M., Rayner, P. J., and Wang, Y. P.: Inversion of diurnally varying synthetic CO$_2$: Network optimization for an Australian test case. Global Biogeochem. Cy., 18, GB1044, doi:10.1029/2003GB002136, 2004.

Li, C., Narayanan, V., and Harriss, R. C.: Model estimates of nitrous oxide emissions from agricultural lands in the United States, Global Biogeochem. Cy., 10(2), 297–306, 1996.

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Davis, K. J. and Grainger, C. A.: A near-field tool for simulating the upstream influence of atmospheric observations: The

Stochastic Time-Inverted Lagrangian Transport (STILT) model. J. Geophys. Res., 108, (D16), 4493, doi:10.1029/2002JD003161, 2003

Lin, J. C., Gerbig, C., Wofsy, S. C., Andrews, A. E., Daube, B. C., Grainger, C. A., Stephens, B. B., Bawkin, P. S., and Hollinger, D. Y.: Measuring fluxes of trace gases at regional scales by Lagrangian observations: Application to the CO$_2$ Budget and Rectification Airborne (COBRA) study, J. Geophys. Res., 109, D15304, doi:10.1029/2004JD004754, 2004.

Lin, J. C., Gerbig, C., Wofsy, S. C., Chow, V. Y., Gottlieb, E. W., Daube, B. C. and Matross, D. M.: Designing Lagrangian experiments to measure regional-scale trace gas fluxes, J. Geophys. Res., 112, D13312, doi:10.1029/2006JD008077, 2007.

Macatangay, R., Warneke, T., Gerbig, C., Körner, S., Ahmadov, R., Heimann, M., and Notholt, J.: A framework for comparing remotely sensed and in-situ CO$_2$ concentrations, Atmos. Chem. Phys., 8, 2555–2568, doi:10.5194/acp-8-2555-2008, 2008.

Mahecha, M. D., Reichstein, M., Lange, H., Carvalhais, N., Bernhofer, C., Grünwald, T., Papale, D. and Seufert, G.: Characterizing ecosystem-atmosphere interactions from short to interannual time scales, Biogeosciences, 4, 743-758, 2007, http://www.biogeosciences.net/4/743/2007/.

Manning, A. J., Ryall, D. B., and Derwent, R. G.: Estimating European emissions of ozone-depleting and greenhouse gases using observations and a modeling back-attribution technique. J. Geophys. Res., 108(D14), 4405, doi:10.1029/2002JD002312, 2003

Matross, D. M., Andrews, A. E., Pathmathevan, M., Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Lee, J. T., Zhao, C., Bawkin, P., Munger, J. W., and Hollinger, D. Y.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, Tellus, 58B, 344–358, 2006.

Messager, C., Schmidt, M., Ramonet, M., Bousquet, P., Simmonds, P., Manning, A., Kazan, V., Spain, G., Jennings, S. G. and Ciais, P.: Ten years of CO$_2$, CH$_4$, CO and N$_2$O fluxes over Western Europe inferred from atmospheric measurements at Mace Head, Ireland, Atmos. Chem. Phys. Discuss., 8, 1191–1237, doi:10.5194/acpd-8-1191-2008, 2008.

Michalak, A. M., Bruhwiler, L. M., and Tans, P. P.: A geostatistical approach to surface flux estimation of atmospheric trace gases. J. Geophys. Res., 109, D14109, doi:10.1029/2003JD004433, 2004.

Miller, S. M., Matross, D. M., Andrews, A. E., Millet, D. B., Longo, M., Gottlieb, E. W., Hirsch, A., Gerbig, C., Lin, J. C., Daube, B. C., Hudman, R. C., Dias, P. L. S., Chow, V. Y., and Wofsy, S. C.: Sources of carbon monoxide and formaldehyde in North America determined from high-resolution atmospheric data, Atmos. Chem. Phys., 8, 7673–11451, doi:10.5194/acp-8-7673-2008, 2008.

Nevison, C. and Holland, E.: A reexamination of the impact of anthropogenically fixed nitrogen on atmospheric N$_2$O and the stratospheric O-3 layer, J. Geophys. Res., 102(D21), 25519–25536, 1997.

Nevison, C. D., Lueker, T. J., and Weiss, R. F.: Quantifying the nitrous oxide source from coastal upwelling, Global Biogeochem. Cy., 18, GB1018, doi:10.1029/2003GB002110, 2004.

Olivier, J. G. J., Bouwman, A. F., Van der Hoek, K. W., and Berdowski, J. J. M.: Global air emission inventories for anthropogenic sources of NO$_x$, NH$_3$ and N$_2$O in 1990, Environ. Pollut., 102, 135–148, 1998.

Olivier, J. G. J., Van Aardenne, J. A., Dentener, F., Pagliari, V., Ganzeveld, L. N. and Peters, J. A. H. W.: Recent trends in global greenhouse gas emissions: regional trends 1970-2000 and spatial distribution of key sources in 2000, Environ. Sci., 81–99, doi:10.1080/15693430500400345, 2005.

Peylin, P.: Influence of transport uncertainty on annual mean and seasonal inversions of atmospheric CO$_2$ data, J. Geophys. Res., 107(D19), 4385, doi:10.1029/2001JD000857, 2002.

Peylin, P., Bousquet, P., Ciais, P., and Monfray, P.: Time-Dependant vs Time-Independant inversion of the atmospheric CO$_2$ observations: consequences for the regional fluxes, Inverse Methods in Global Biogeochemical Cycles, edited by: Kasibhatla, P., Heimann, M., Rayner, R., Mahowald, N., Prinn, R. G., and Hartley, D. E., American Geophysical Union, 114, 324 pp., 1999.

Peylin, P., Rayner, P. J., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P. and contributors, A.: Daily CO$_2$ flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, Atmos. Chem. Phys., 5, 3173–3186, doi:10.5194/acp-5-3173-2005, 2005.

Popa, E.: Continuous tall tower multi-species measurements in Europe for quantifying and understanding land-atmosphere carbon exchange, Friedrich Schiller University, PhD, Jena, Germany, 2008.

Press, W., Teukolsky, S., Vetterling, W., and Flannery, B.: Numerical Recipes, Cambridge University Press, Cambridge, 1992.

Prinn, R., Cunnold, D., Rasmussen, R., Simmonds, P., Alyea, F., Crawford, A., Fraser, P., and Rosen, R.: Atmospheric emissions and trends of nitrous oxide deduced from 10 years of ALE-GAGE data, J. Geophys. Res., 95(D11), 18369–18385, 1990.

Ramaswamy, V.: Radiative forcing of climate change, Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK, 349–416. 2001.

Ravishankara, A. R., Daniel, J. S., and Portmann, R. W.: Nitrous Oxide (N$_2$O): The Dominant Ozone-Depleting Substance Emitted in the 21st Century, Science, 326(5949), 123–125, 2009.

Rodgers, C. D.: Inverse methods for atmospheric sounding: theory and practice, World Scientific, Singapore, 2000.

Schmidt, M., Glatzel-Mattheier, H., Sartorius, H., Worthy, D. E. and Levin, I.: Western European N2O emissions: A top-down approach based on atmospheric observations, J. Geophys. Res., 106(D6), 5507–5516, 2001.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Society for Industrial and Applied Mathematics, University City Science Center, Philadelphia, USA, 2005.

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

Trampert, J. and Snieder, R.: Model Estimations Biased by Truncated Expansions: Possible Artifacts in Seismic Tomography, Science, 271, 1257–1260, 1996.

Uliasz, M. and Pielke, R. A.: Receptor-oriented Lagrangian-Eulerian model of mesoscale air pollution dispersion, Computer Techniques in Environmental Studies, edited by: Zannetti, P., Billerica, MA, USA, 57–68. 1990.

Wang, Y. P. and Barrett, D. J.: Estimating regional terrestrial carbon fluxes for the Australian continent using a multiple-constraint approach I. Using remotely sensed data and ecological observations of net primary production, Tellus, 55B, 270–289, 2003.

Worthy, D. E. J., Platt, A., Kessler, R., Ernst, M., and Racki, S.: The Greenhouse Gases Measurement Program: Measurement Procedures and Data Quality, GGML, 2003.