



Joint CO₂ state and flux estimation with the 4D-Var system EURAD-IM

Johannes Klimpt^{1,2}, Elmar Friese¹, and Hendrik Elbern^{1,2}

¹Rhenish Institute for Environmental Research at the University of Cologne, 50931 Köln, Germany

²Institute for Energy and Climate Research (IEK-8), Forschungszentrum Jülich, 52425 Jülich, Germany

Correspondence to: J.Klimpt, j.klimpt@uni-koeln.de

Abstract. This idealized regional atmospheric inversion study assesses the potential of the 4-dimensional variational (4D-Var) method to estimate CO₂ fluxes and the atmospheric CO₂ concentration state jointly. In order to distinguish and quantify the surface-atmosphere CO₂ fluxes, combining anthropogenic CO₂ emissions, photosynthesis, and respiration, we include uncertainties of initial values, which arise from highly uncertain surface fluxes and night-time transport. Therefor a new calculation of the background error standard deviation for the CO₂ fluxes was developed. To suppress spurious wiggles occurring from advection, an absolute monotone advection scheme with low numeric diffusion and its adjoint has been implemented. The inversion by the EUROpean Air pollution Dispersion-Inverse Model (EURAD-IM) with 5 km resolution in Central Europe is validated by synthetic half hourly measurements from eleven concentration towers. A significant improvement of the analysis is shown if initial values and CO₂ fluxes are optimised jointly, compared to optimising CO₂ fluxes alone, without estimating uncertainty of atmospheric concentration. We find that joint estimation of carbon fluxes and initial states requires a careful balance of the background error covariance matrices but enables a more detailed analysis of atmospheric CO₂ and the surface-atmosphere fluxes.

1 Introduction

The quantification of terrestrial CO₂ sinks and sources is important in order to understand the carbon cycle and its fundamental role in the global climate system. On a global scale, approximately 30% of the anthropogenic emissions are removed by oceans (Wanninkhof et al., 2013) as well as 30% by terrestrial biosphere (Sitch et al., 2013), the latter one showing much larger temporal and spatial variability (Jung et al., 2011).

One widely used method for the estimation of the surface-atmosphere CO₂ exchange is atmospheric inverse modelling, using CO₂ concentration measurements and atmospheric transport models. On a global scale atmospheric inversions improve the quantification of natural and anthropogenic carbon fluxes (Tans et al., 1990; Enting et al., 1993; Rödenbeck et al., 2003; Rayner et al., 2008), although the uncertainty between studies is large, especially between continents (Gurney et al., 2002; Baker et al., 2006; Peters et al., 2010).

Locally the increase of spatio-temporal model resolution leads to even higher uncertainties of atmospheric inversion studies (Lauvaux et al., 2008; Broquet et al., 2011), mainly for two reasons: (1) A derivation of appropriate a priori surface carbon



fluxes is difficult due to spatial heterogeneity. In particular European land use is highly variable since large parts of the terrestrial biosphere are actively managed by humans and are often embedded in densely populated areas, including cities and industrial areas (Peters et al., 2010). (2) Atmospheric transport modelling introduces uncertainty to atmospheric CO₂ mixing ratio of several ppm due to advection (Lin and Gerbig, 2005) and vertical mixing (Gerbig et al., 2008). Although uncertainty from transport model errors is considerably lower than uncertainty from CO₂ surface fluxes (Tolk et al., 2009), the skills of the transport model are essential for CO₂ inversions.

Additionally to the mentioned difficulties of proper transport models and surface-atmosphere CO₂ fluxes one major obstacle for CO₂ inversions is the sparsity of representative concentration observations, e.g. Wu et al. (2016) addresses the question of optimal location of observations in inversion systems. To address uncertainty appropriately to the mentioned model processes and input data is among the most challenging issues for atmospheric CO₂ inversions.

Tolk et al. (2011) compare the uncertainty reduction for different strategies of inverse carbon modelling, suggesting an optimisation of biosphere model parameters or a pixel inversion of respiration and photosynthesis on a regional scale.

To reduce uncertainty of a priori fluxes, additional carbon flux measurements from eddy covariance towers can be used. Lauvaux et al. (2008), Cooley et al. (2013), and Zhu et al. (2014) combine this so-called bottom-up approach with top-down atmospheric inversion in order to quantify regional carbon fluxes.

To reduce errors in atmospheric transport processes, different strategies are investigated. For example additional information from radon measurements (Broquet et al., 2011) or the energy balance (Tolk et al., 2009) are used to gain information of atmospheric CO₂ dispersion. Studies with an horizontal resolution of a few km (Ahmadov et al., 2007, 2009; Pillai et al., 2011) increase confidence in modelled CO₂ mixing ratios compared to coarser resolution, although comparisons of meso-scale models reveal discrepancies for the vertical boundary layer (Sarrat et al., 2007; Geels et al., 2007). Uncertainty of vertical model transport is considerably higher during night due to uncertain mixing height (Lauvaux et al., 2008; Dolman et al., 2009; Kretschmer et al., 2014), leading to a poor representativity of concentration during morning hours. Still, optimisation of initial conditions is often mitigated by long spin-up periods or is investigated at a coarse subspace (Peylin et al., 2005; Lauvaux et al., 2008).

This study investigates how CO₂ flux uncertainty from inversions can be reduced by a proper accounting of uncertainty of atmospheric concentration. The feasibility and limitations of a comprehensive 4D-Var inversion system to analyse anthropogenic emissions, photosynthesis, ecosystem respiration, and initial values at each grid cell jointly, are validated in this study.

This paper is organised as follows. Section 2 delivers the theoretical background of the joint optimisation of fluxes and initial values, Sect. 3 describes the implementation into the EURAD-IM. Numerical experiments validating the impact of erroneous initial values are presented and discussed in Sect. 4. Section 5 summarises the results.

2 Theory for joint variational assimilation

In this study we want to optimise two parameters with the 4D-Var data assimilation method: initial values and flux factors jointly. While the proceeding of optimising initial values is well-known, we want to visualise the basic concept of optimising

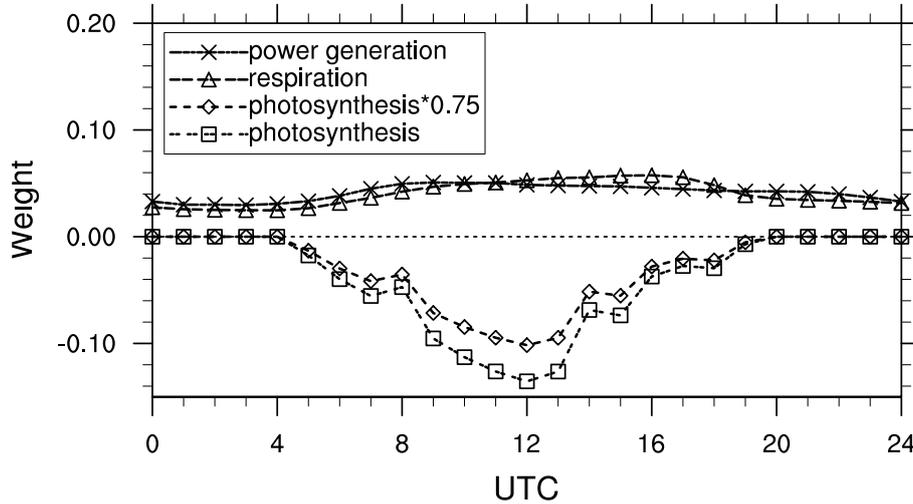


Figure 1. The diurnal time profile of power generation given by Memmesheimer et al. (1995) and spatial averaged biogenic fluxes calculated with WRF-CLM at 24 July 2012 (absolute weights sum up to 1). An example with an analysed flux factor of 0.75 for photosynthesis is shown, decreasing the total amount but preserving the profile.

flux factors and joint optimisation. The main idea (Elbern et al., 2000) is to reduce the degree of freedom of the flux rate space by not optimising the fluxes at each time step t_i . Rather it is pointed out that, due to the better knowledge of the diurnal cycle of fluxes, one efficient parameter to optimise is their diurnal amplitude, as simulated by the forward model.

This study optimises three fluxes, anthropogenic emissions, photosynthesis, and respiration. For simplicity we will define the flux space to be of the same dimension as the state space (\mathbb{R}^n). The flux rate vector $\mathbf{f} \in \mathbb{R}^n$ scales the a priori knowledge of the background flux \mathbf{U}^b by an optimisation factor per grid point and per flux type, while the relative diurnal profile variation remains unchanged. For notational convenience we define $\mathbf{U}^b \in \mathbb{R}^{n \times n}$ to be a diagonal matrix.

The actual fluxes are thus ($\text{diag}(\mathbf{a})$ designates a diagonal matrix with entries of the vector \mathbf{a})

$$\mathbf{U}_{i+1/2} = \text{diag}(\mathbf{U}_{i+1/2}^b \mathbf{f}) \in \mathbb{R}^{n \times n}, \quad \forall i = 0, \dots, N-1, \quad (1)$$

with $\mathbf{U}_{i+1/2}$ denoting the fluxes in the time interval $[t_i, t_{i+1}]$ introduced instantaneously at $t_{i+1/2}$. For notational convenience subindices in this paper will be used exclusively to identify discrete time. The simulated time period is defined as $[t_0, t_1, \dots, t_N]$, $(\cdot)_i$ refers to t_i , $(\cdot)_{i+1/2}$ to $(t_i + t_{i+1})/2$, and $(\cdot)_{i,j}$ to the interval $[t_i, t_j]$.

Figure 1 shows the diurnal cycle of the three CO_2 fluxes, anthropogenic emissions, photosynthesis, and respiration, where the photosynthesis is exemplarily rescaled. The actual parameter of optimisation will be $\mathbf{g} := \ln(\mathbf{f})$ due to two reasons: (1) The transformation results in analog partial costs for flux factors \mathbf{g} and $1/\mathbf{g}$. (2) Since $\mathbf{f} > \mathbf{0}$ it can be described by a log-normal distribution, resulting in a Gaussian probability density function for \mathbf{g} . Fletcher and Zupanski (2006) and Fletcher (2010) describe 4D-Var systems with hybrid Gaussian and log-normal distribution in general.

The idea for the main advantage of joint optimisation of initial values and flux factors is depicted in Fig. 2. While the influence

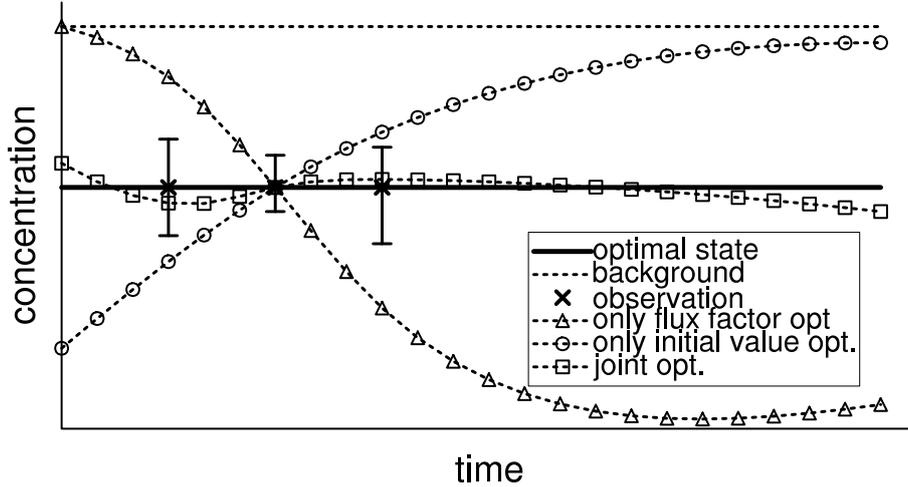


Figure 2. If a priori initial values and flux rates are defective, the joint optimisation combines the possibility of analysing both error sources. Optimising only initial values leads to strong forcing at the start of the assimilation window. Optimising only flux rates can not change the initial values and may force the analysis to give more weight at the starting time than desirable.

of optimised initial values decreases with longer forecast period, optimised flux factors can hardly adjust a mismatch between model and observation at the beginning of a forecast.

In the following we briefly present the theoretical background and the formulation of the cost function \mathcal{J} , following Elbern et al. (2000) and Elbern et al. (2007). We assume a background field $\mathbf{x}_0^b \in \mathbb{R}^n$, which is the initial concentration value of the passive tracer CO_2 , and $\mathbf{g}^b = (0, \dots, 0) \in \mathbb{R}^n$, the background logarithmic flux factors, are given. Our a priori knowledge about the fluxes is actually represented by the flux strength $\mathbf{U}_{t_{1/2}, \dots, t_{N-1/2}}^b$ given in $\frac{\mu\text{molCO}_2}{\text{m}^2\text{s}}$. If $[t_0, \dots, t_N]$ denotes the assimilation window considered, we set further observations $\mathbf{y}_i \in \mathbb{R}^{p_i}$ at any time step t_i .

In the following subscripts i always refer to time t_i . The background field \mathbf{x}_0^b evolves from time t_0 to t_i by the use of the model $\mathcal{M}_{0,i}$:

$$10 \quad (\mathbf{x}_i, \mathbf{f}) := \mathcal{M}_{0,i}(\mathbf{x}_0, \mathbf{f}). \quad (2)$$

The observation operator \mathcal{H}_i projects the model state and the flux factor from \mathbb{R}^{2n} to the observation space \mathbb{R}^{p_i} . This mapping \mathcal{H}_i allows to compare observation \mathbf{y}_i with the model equivalent $\mathcal{H}_i(\mathbf{x}_i^b, \mathbf{g}^b)$, by the innovation vector \mathbf{d}_i

$$\mathbf{d}_i := \mathbf{y}_i - \mathcal{H}_i(\mathbf{x}_i^b, \mathbf{g}^b) = \mathbf{y}_i - \mathcal{H}_i \mathcal{M}_{0,i}(\mathbf{x}_0^b, \mathbf{g}^b). \quad (3)$$

Following Courtier et al. (1994) we define the cost function \mathcal{J} in dependence of the increments, which is the difference between the optimisation parameters and the a priori knowledge of initial values and flux factors, $\delta \mathbf{x}_0 := \mathbf{x}_0 - \mathbf{x}_0^b$ and $\delta \mathbf{g} := \mathbf{g} - \mathbf{g}^b$. The mutually uncorrelated error covariance matrices for the initial values (IV), flux factors (FF), and observations (O) are called



$\mathbf{B} \in \mathbb{R}^{n \times n}$, $\mathbf{K} \in \mathbb{R}^{n \times n}$, and $\mathbf{R}_i \in \mathbb{R}^{p_i \times p_i}$ respectively. The cost function \mathcal{J} can now be defined as

$$\begin{aligned} \mathcal{J}(\delta \mathbf{x}_0, \delta \mathbf{g}) &:= \mathcal{J}^{\text{IV}} + \mathcal{J}^{\text{FF}} + \mathcal{J}^{\text{O}} \\ &:= \frac{1}{2} (\delta \mathbf{x}_0)^{\text{T}} \mathbf{B}^{-1} (\delta \mathbf{x}_0) + \frac{1}{2} (\delta \mathbf{g})^{\text{T}} \mathbf{K}^{-1} (\delta \mathbf{g}) + \\ &\quad \frac{1}{2} \sum_{i=0}^N [\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} (\delta \mathbf{x}_0^{\text{T}}, \delta \mathbf{g}^{\text{T}})^{\text{T}}]^{\text{T}} \mathbf{R}_i^{-1} [\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} (\delta \mathbf{x}_0^{\text{T}}, \delta \mathbf{g}^{\text{T}})^{\text{T}}]. \end{aligned} \quad (4)$$

The operators $\mathbf{H}_i \in \mathbb{R}^{2n \times p_i}$, $\mathbf{M}_{0,i} \in \mathbb{R}^{2n \times 2n}$ are tangent linear approximations with

$$\mathcal{H}_i(\mathbf{x}_i + \delta \mathbf{x}_i, \mathbf{g} + \delta \mathbf{g}) \approx \mathcal{H}_i(\mathbf{x}_i, \mathbf{g}) + \mathbf{H}_i (\delta \mathbf{x}_i^{\text{T}}, \delta \mathbf{g}^{\text{T}})^{\text{T}}, \quad (5)$$

$$\mathcal{M}_{0,i}(\mathbf{x}_i + \delta \mathbf{x}_0, \mathbf{g} + \delta \mathbf{g}) \approx \mathcal{M}_{0,i}(\mathbf{x}_0, \mathbf{g}) + \mathbf{M}_{0,i} (\delta \mathbf{x}_0^{\text{T}}, \delta \mathbf{g}^{\text{T}})^{\text{T}} \quad (6)$$

for sufficiently small $\delta \mathbf{x}_0$ and $\delta \mathbf{g}$. Consequently the gradient of \mathcal{J} is

$$\nabla \mathcal{J}(\delta \mathbf{x}_0, \delta \mathbf{g}) = \begin{pmatrix} \mathbf{B}^{-1} \delta \mathbf{x}_0 \\ \mathbf{K}^{-1} \delta \mathbf{g} \end{pmatrix} - \sum_{i=0}^N (\mathbf{M}_{0,i})^{\text{T}} \mathbf{H}_i^{\text{T}} \mathbf{R}_i^{-1} \begin{bmatrix} \mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} \begin{pmatrix} \delta \mathbf{x}_0 \\ \delta \mathbf{g} \end{pmatrix} \end{bmatrix}, \quad (7)$$

where $(\mathbf{M}_{0,i})^{\text{T}}$ and \mathbf{H}_i^{T} are the adjoint operators, with $(\mathbf{M}_{0,i})^{\text{T}}$ modelling backwards in time from t_i to t_0 . Since the construction of the adjoint model uses the derivative with respect to the optimisation parameter we have to distinguish between the adjoint model related to initial values \mathbf{x}_0 and flux factors \mathbf{g} . See Appendix A for a detailed derivation of $(\mathbf{M}_{0,i})^{\text{T}}$ with respect to the initial values and flux factors. A careful balancing of the background error covariance matrices (BECM) \mathbf{B} and \mathbf{K} and the matrix \mathbf{R} for the observation errors is decisive for the analysis.

3 Model description

The model applied here is derived from the multiscale EUROpean Air pollution Dispersion-Inverse Model (EURAD-IM), which is a state of the art Chemistry Transport Model (CTM) (Elbern et al., 2007; Goris and Elbern, 2015) including gas phase, aerosols and pollen.

- 20 The model domain uses the Lambert conformal projection, where the center of the coarse domain is located at 54° N latitude, 12.5° E longitude and has the staggered Arakawa C-grid structure in this study. A nesting configuration with a horizontal resolution of 15×15 km² and 5×5 km² (Fig. 3 (a)) is used. The model domain includes 23 vertical levels using terrain following σ coordinates up to 100 hPa. For the parent domain constant lateral boundary conditions are taken, the usage of lateral boundary values from global as provided in Copernicus Atmosphere Monitoring Service (CAMS) is feasible as well.
- 25 For simplicity the background concentration of CO₂ for all domains is assumed to be constant 386 ppmV.

3.1 Observations

The measurement towers used in this study are part of the FLUXNET (<http://fluxnet.ornl.gov/>) network, with a spatial focus on the TR32 Rur catchment area (<http://tr32new.uni-koeln.de/>) in western Germany. All stations are listed in detail in Table

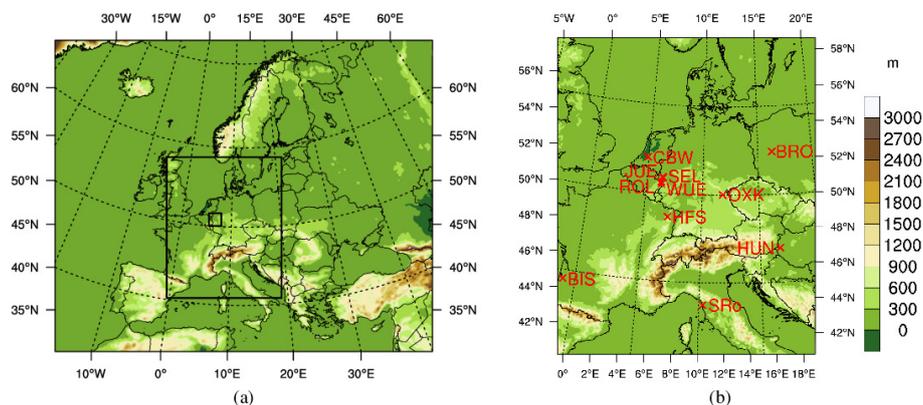


Figure 3. Topography of (a) the nests with 15 and 5 km resolution and (b) the 5 km domain with the location of the synthetic observation sites, see also Table 1.

Table 1. Stations from FLUXNET (<http://fluxnet.ornl.gov/>) used in this study. ENF = Evergreen Needleleaf Forests, DBF = Deciduous Broadleaf Forests

Code	Lon, °	Lat, °	Elevation, m asl	Measurement Height, m agl	Name	Vegetation
BIS	-1.23	52.4342	120	42	Biscarosse, Fra	ENF
BRO	16.2995	52.4342	83	5	Brody, Pol	Cropland
CBW	4.9270	51.9710	5	20, 60, 120, 200	Cabauw, Ned	Grassland
HUN	16.6520	46.9559	234	10, 48, 82, 117	Hegyhatsal, Hun	Cropland
HFS	7.0655	48.6742	304	22	Hesse Forest Sarre., Fra	DBF
JUE	6.4094	50.9100	101	20, 100	Juelich, Deu	-
OXK	11.8083	50.0301	1020	23, 90, 163	Ochsenkopf, Deu	ENF
ROL	6.3041	50.6219	486	3	Rollesbroich, Deu	Grassland
SEL	6.4497	50.8706	109	3	Selhausen, Deu	Cropland
SRo	10.2909	43.7320	6	23.5	San Rossore 2, Ita	ENF
WUE	6.3310	50.5049	578	38	Wuestebach, Deu	ENF

1. Measurements are taken every 30 minutes and the data will be assimilated between 06 and 18 UTC, resulting in 340 observations in total. Except from two towers (BIS and WUE) all towers provide measurements below 30 m, which is in the surface layer of our model. Four towers provide also measurements in several vertical heights. To ease comparison no distinction between high altitude stations as Ochsenkopf and flat stations is made here.



3.2 Meteorological Input

- 5 The EURAD-IM uses the Weather Research and Forecasting (WRF) model version 3.6.1 (Skamarock et al., 2008), a mesoscale numerical weather prediction system as meteorological driver, to provide hourly meteorological input. The necessary meteorological initial and lateral boundary data are taken from six-hourly ECMWF (<http://www.ecmwf.int>) reanalysis, interpolated to $0.225^\circ \times 0.225^\circ$. The most important variables are wind, temperature, shortwave radiation and planetary boundary layer height.

5 3.3 Anthropogenic CO₂ emissions

- Anthropogenic emissions, introduced as first guess for inversion, are taken from Toegepast Natuurwetenschappelijk Onderzoek (TNO) inventory (H. Denier van der Gon, personal communication, May 2013), following the methodology outlined in Kuenen et al. (2014) and Pouliot et al. (2012) for air pollutants. It provides point and area sources of CO₂ splitted into ten source categories for Europe for the reference year 2005. Annual values are disaggregated for each source category in dependence of month, day and hour. See also Memmesheimer et al. (1995) for details. Hourly values are then interpolated to model time step. Area sources are scaled down to 5 km resolution, using land cover and infrastructure information with the Geographic Information System (GIS).

3.4 Biogenic CO₂ fluxes

- Hourly photosynthesis, leaf-, and soil-respiration are calculated using the Community Land Model (CLM) version 4.0 (Oleson et al., 2010), which is used as land surface scheme for WRF. While photosynthesis is already included in this configuration of CLM according to Collatz et al. (1991) and Farquhar et al. (1980), leaf respiration was added according to the same publications. While leaf respiration is in dependence of the maximum carboxylation capacity, the calculation of photosynthesis depends further on the internal leaf CO₂ and O₂ pressure, the amount of radiation, the availability of the enzyme Rubisco, and the Michaelis-Menten constants for CO₂ and O₂. The Leaf Area Index (LAI) is monthly Moderate Resolution Imaging Spectroradiometer (MODIS) data, averaged for the years 2001-2010, provided as a part of the WRF preprocessing system. The CLM sub-grid hierarchy is simplified to allow only one Plant Functional Type (PFT) per grid cell, following Oleson et al. (2010) (see Fig. 1.1 in that publication for details).

Soil respiration is implemented as an Arrhenius type equation in dependence of temperature of soil layer four (ca. 11-15 cm depth), following Lloyd and Taylor (1994, Eq. (8)). Other CO₂ fluxes from the biosphere are neglected in this work.

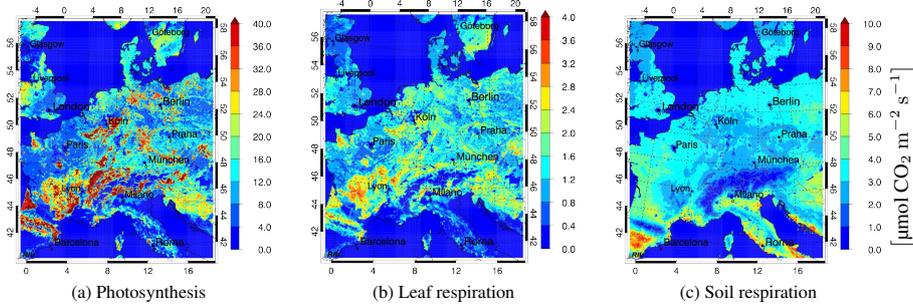


Figure 4. WRF-CLM output of photosynthesis, leaf and soil respiration at 23 July 2012 13 UTC with 5 km horizontal resolution.

3.5 Forward model of EURAD-IM

- 5 The transport-diffusion equation which describes the dispersion of the mixing ratio of CO_2 , $C(\mathbf{r}, t)$ given in dependence of spatial location $\mathbf{r} \in \mathbb{R}^3$ and time $t \in [t_0, t_N]$, as a passive tracer can be written as (Elbern and Schmidt, 2001)

$$\frac{\partial C}{\partial t} + \underbrace{(\mathbf{u} \cdot \nabla)C}_{\text{advection}} - \underbrace{\frac{1}{\rho} \nabla \cdot (\rho \mathbf{K}^e \nabla C)}_{\text{turbulent diffusion}} = \underbrace{e}_{\text{flux}}. \quad (8)$$

The equation includes terms for advection, turbulent diffusion, and additional fluxes. Here \mathbf{u} designates the wind field, ρ the air density and \mathbf{K}^e is the eddy diffusion tensor. The sinks and sources to the atmosphere, consisting of anthropogenic emissions, biogenic respiration and photosynthesis are aggregated to the flux e .

- The EURAD-IM forward model discretises Eq. (8) with an operator splitting approach due to different numerical characteristics of the major transport processes advection and turbulent diffusion (Yanenko, 1971). The operators for advection and diffusion are processed for time steps t_i to $t_{i+1/2}$ and $t_{i+1/2}$ to t_{i+1} while fluxes are inserted instantaneously at time points $t_{i+1/2}$. A symmetric order of the operators $\mathbf{A}_{i,i+1/2}^1 \in \mathbb{R}^n$ (advection in sequence x-y-z), $\mathbf{A}_{i+1/2,i+1}^2 \in \mathbb{R}^n$ (advection in sequence z-y-x), $\mathbf{D}_{i,i+1/2}, \mathbf{D}_{i+1/2,i+1} \in \mathbb{R}^n$ (turbulent diffusion), and $\mathbf{F}_{i+1/2}^f \in \mathbb{R}^n$ (fluxes) is chosen:

$$\mathbf{x}_{i+1} = \left(\mathbf{A}_{i+1/2,i+1}^2 \mathbf{D}_{i+1/2,i+1} \mathbf{F}_{i+1/2}^f \mathbf{D}_{i,i+1/2} \mathbf{A}_{i,i+1/2}^1 \right) \mathbf{x}_i, \quad (9)$$

where the flux operator is defined as (background fluxes \mathbf{U}^b are given in $\frac{\mu\text{molCO}_2}{\text{m}^2\text{s}}$ and $\Delta t := t_{i+1} - t_i$ is given in seconds as well)

$$\mathbf{F}_{i+1/2}^f (\mathbf{D}_{i,i+1/2} \mathbf{A}_{i,i+1/2}^1 \mathbf{x}_i) := (\mathbf{D}_{i,i+1/2} \mathbf{A}_{i,i+1/2}^1 \mathbf{x}_i) + \Delta t \mathbf{U}_{i+1/2}^b \mathbf{f}. \quad (10)$$

We remember that $\mathbf{U}_{i+1/2}^b \mathbf{f}$ represents actually three fluxes, each with an independent flux factor for anthropogenic emissions, photosynthesis, and respiration respectively.

Tests with a non-monotone advection scheme display strong deterioration of CO_2 mixing ratio if strong spatial gradients of the wind strength occur, see Fig. 5. To account for this problem the Walcek scheme (Walcek, 2000) is used to solve the advection

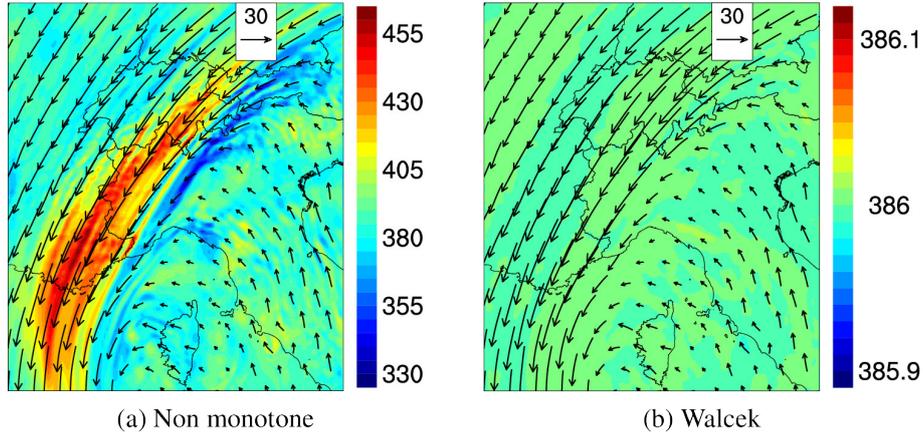


Figure 5. Zoom of the 5 km horizontal resolution grid, started with constant mixing ratio after four hour run-time at a high vertical level ($\sim 373 - 282$ hPa) barely influenced from surface fluxes. The wind field is normalised to 30 m/s.

- 5 process. Applying dimensional dependent densities, allows splitting of the forward equation into one-dimensional realisations. The Walcek scheme solves the 1-dim transport equation

$$\frac{\partial C}{\partial t} = - \frac{\partial(u \cdot C)}{\partial x} \quad (11)$$

based on the van Leer algorithm (Van Leer, 1977) with monotonic constraints and a flux adjustment around local mixing ratio extrema. This aggregates mass numerically and reduces numerical diffusion at
 10 extrema. It is absolute monotone, positive definite, of order two, produces only low numeric diffusion, and is numerical efficient. Although an exact parallelisation of the Walcek scheme is numerically not efficient, a good approximation with a small overlap of two grid cells was realised, which preserves the numerical efficiency.

3.6 Adjoint EURAD-IM model

- 15 The adjoint model \mathbf{M}^T discretises the adjoint of the forward PDE, Eq. (8). Let C^* designate the adjoint variable, then the adjoint PDE has the form

$$- \frac{\partial C^*}{\partial t} - \nabla(\mathbf{u}C^*) - \nabla \left(\rho \mathbf{K}^e \nabla \frac{C^*}{\rho} \right) = \phi, \quad (12)$$

where ϕ represents perturbation from the reference CO_2 flux e of Eq. (8).

- Although the flux factor $\mathbf{f} = e^g$ is constant during the forward run, the adjoint flux factor \mathbf{f}^* changes with $(\mathbf{M})^T$ during the
 20 background run and, in contrast to Eq. (9), we need to calculate \mathbf{f}^* as well. The EURAD-IM backward model applies again an



operator splitting to solve Eq. (12), :

$$(\mathbf{M}_{i,i+1})^T \begin{pmatrix} \mathbf{x}_{i+1}^* \\ \mathbf{f}_{i+1}^* \end{pmatrix} = \begin{pmatrix} (\mathbf{T}_{i,i+1})^T \mathbf{x}_{i+1}^* \\ \Delta t \mathbf{U}_{i+1/2}^b (\mathbf{T}_{i+1/2,i+1})^T \mathbf{x}_{i+1}^* + \mathbf{f}_{i+1}^* \end{pmatrix} \quad (13)$$

with $\mathbf{T}_{i,i+1}$ merging the transport operators from time t_i to t_{i+1}

$$\mathbf{T}_{i,i+1} := \mathbf{A}_{i+1/2,i+1}^2 \mathbf{D}_{i+1/2,i+1} \mathbf{D}_{i,i+1/2} \mathbf{A}_{i,i+1/2}^1, \quad \mathbf{T}_{i+1/2,i+1} = \mathbf{A}_{i+1/2,i+1}^2 \mathbf{D}_{i+1/2,i+1}. \quad (14)$$

25 See Appendix A for a detailed derivation of Eq. (13).

The adjoint diffusion and flux operators \mathbf{D}^T and $(\mathbf{F}^f)^T$ are constructed with the automatic differentiation tool Tapenade (Hascoet and Pascual, 2013) from the discretisation of the forward PDE Eq. (8): The derivative of the functions specified by \mathbf{D} and \mathbf{F}^f are calculated applying the chain rule to get the tangent linear of the forward operators. This linearised code is transposed afterwards resulting in \mathbf{D}^T and $(\mathbf{F}^f)^T$.

To construct the adjoint of the advection operators $(\mathbf{A}^1)^T$ and $(\mathbf{A}^2)^T$ a different approach has been chosen. The Walcek scheme is not differentiable due to the monotony and automatic differentiation is not directly applicable to obtain adjoint code. Following the suggestion of Gou and Sandu (2011) the continuous adjoint for advection has been used. The adjoint PDE of Eq. (11)

$$-\frac{\partial C^*}{\partial t} = \frac{\partial (u \cdot C^*)}{\partial x}, \quad (15)$$

can be also used by the forward Walcek scheme, if the reverse winds and a rescaling for C^* is used. The work cited above also suggests for 4D-Var optimisation problems to construct adjoint routines for diffusion and advection, as it is done here: automatic differentiation for diffusion and the continuous adjoint for advection.

3.7 Preconditioning of the cost function and minimisation

The background error covariance matrices (BECM) \mathbf{B} and \mathbf{K} are designed to capture the spatial correlation between initial values and surface fluxes respectively. Increasing radii of influence lead to larger condition numbers of \mathbf{B} , which makes efficient minimisation more difficult. Using the increments $\delta \mathbf{x}_0$ and $\delta \mathbf{g}$ multiplied with the inverse square roots of \mathbf{B} and \mathbf{K} , a formulation of \mathcal{J} equivalent to Eq. (4) is possible:

$$\mathbf{v} := \mathbf{B}^{-1/2} \delta \mathbf{x}_0 = \mathbf{B}^{-1/2} (\mathbf{x}_0 - \mathbf{x}_0^b) \quad \text{and} \quad \mathbf{w} := \mathbf{K}^{-1/2} \delta \mathbf{g} = \mathbf{K}^{-1/2} (\mathbf{g} - \mathbf{g}^b),$$

$$\mathcal{J}(\mathbf{v}, \mathbf{w}) := \frac{1}{2} \mathbf{v}^T \mathbf{v} + \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{1}{2} \sum_{i=0}^N [\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} (\mathbf{v}^T, \mathbf{w}^T)^T]^T \mathbf{R}_i^{-1} [\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} (\mathbf{v}^T, \mathbf{w}^T)^T]. \quad (16)$$

The gradient of \mathcal{J} with respect to \mathbf{v} and \mathbf{w} reads then

$$\nabla \mathcal{J}(\mathbf{v}, \mathbf{w}) = \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{T/2} & 0 \\ 0 & \mathbf{K}^{T/2} \end{pmatrix} \sum_{i=0}^N (\mathbf{M}_{0,i})^T \mathbf{H}_i^T \mathbf{R}_i^{-1} \begin{bmatrix} \mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} \end{bmatrix}. \quad (17)$$

20 One iteration cycle of the 4D-Var system EURAD-IM to jointly analyse initial values and flux factors consists of:



1. The forward run, solving Eq. (8). Our a priori knowledge $\mathbf{x}_0^b, \mathbf{g}^b$ is used as first guess, allowing to apply only the square root (and its transposed) of the BECM's \mathbf{B} and \mathbf{K} .
2. The calculation of the gradient of \mathcal{J} during the backward time loop and its preconditioning resulting in:

$$\begin{pmatrix} \mathbf{B}^{\text{T}/2} & 0 \\ 0 & \mathbf{K}^{\text{T}/2} \end{pmatrix} \sum_{i=0}^N (\mathbf{M}_{0,i})^{\text{T}} \mathbf{H}_i^{\text{T}} \mathbf{R}_i^{-1} \left[\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} \right].$$
3. The usage of limited memory quasi-Newton minimiser L-BFGS (Liu and Nocedal, 1989; Nocedal, 1980), modified for parallel usage, which calculates a new vector $(\mathbf{v}^{\text{T}}, \mathbf{w}^{\text{T}})^{\text{T}}$ and saves it for the next iteration.
4. The determination of improved initial states $\mathbf{x}_0 = \mathbf{B}^{1/2} \mathbf{v} - \mathbf{x}_0^b$ and flux factors $\mathbf{g} = \mathbf{K}^{1/2} \mathbf{w} - \mathbf{g}^b$ that can now be used iteratively for step 1.
- 5 Steps 1 to 4 are repeated until the convergence criterion is fulfilled. In practice the two most recent evaluations of the cost function must differ by less than $1.0\text{E}-6$.

3.8 Background error covariance matrices modelling

Both BECM's \mathbf{B} and \mathbf{K} are modelled using the diffusion approach. Following Weaver and Courtier (2001) a generalised diffusion equation (GDE) is formulated with a polynomial in the Laplacian, which is self-adjoint. This allows a factorisation of \mathbf{B} into $\mathbf{B}^{1/2} \mathbf{B}^{\text{T}/2}$ and analog for \mathbf{K} . Also anisotropic and inhomogeneous spatial radii of influence can be modelled in this manner.

3.8.1 Initial value background error covariance matrix

The modelling of \mathbf{B} is the same as in Elbern et al. (2007) and is given in short form here for convenience. The background error standard deviation depends on the initial concentration

$$\sqrt{(B_{k,k})} = \mathbf{x}_0^b(k) \cdot c^{\text{IV}}, \quad (18)$$

with $c^{\text{IV}} = 0.025$. The decomposition of \mathbf{B} (and analog for \mathbf{K}) is given by

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma \quad (19)$$

$$= \Sigma \mathbf{C}^{1/2} \mathbf{C}^{\text{T}/2} \Sigma \quad (20)$$

$$= (\Sigma \Lambda \mathbf{L}_v^{1/2} \mathbf{L}_h^{1/2} \mathbf{W}^{-1/2}) (\mathbf{W}^{-1/2} \mathbf{L}_h^{\text{T}/2} \mathbf{L}_v^{\text{T}/2} \Lambda \Sigma) \quad (21)$$

$$= \mathbf{B}^{1/2} \mathbf{B}^{\text{T}/2}, \quad (22)$$

where Σ is the diagonal matrix of standard deviations, \mathbf{C} is the correlation matrix with Λ a diagonal normalisation matrix and $\mathbf{L}_{v,h}$ the vertical and horizontal diffusion operator. \mathbf{W} is also a diagonal matrix accounting for the extension of each vertical layer $k = 1, \dots, k_{\text{max}} = 23$.



3.8.2 Flux factor background error standard deviation

25 We present the variance modelling for the flux factor BECM \mathbf{K} . Joint optimisation of initial values and flux factors requires a distinction of the adjoint models with respect to the two optimisation parameters. Although the flux factor $\mathbf{f} = e^g$ is constant during one iteration, the adjoint flux factor \mathbf{f}^* changes with $(\mathbf{M})^T$ depending on the background flux strength \mathbf{U}^b . We want to investigate the influence of \mathbf{U}^b on $\nabla \mathcal{J}$ hereinafter. Using Eq. (13) iteratively and the transport operator \mathbf{T} (Eq. (14)), we derive

$$(\mathbf{M}_{0,i})^T \begin{pmatrix} \mathbf{x}_i^* \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} (\mathbf{T}_{0,i})^T \mathbf{x}_i^* \\ \sum_{j=1}^i \Delta t \mathbf{U}_{j-1/2}^b (\mathbf{T}_{j-1/2,i})^T \mathbf{x}_i^* \end{pmatrix}, \quad (23)$$

with an equidistant time interval $[t_0, \dots, t_N]$: $\Delta t := t_{i+1} - t_i$ for $i = 0, \dots, N-1$, see Appendix A for a detailed derivation. Thus the adjoint model with respect to the flux factor is proportional to $\mathbf{U}^b (\mathbf{T})^T$.

5 With regard to Eq. (17) the quantity $\mathbf{K}^{T/2} \mathbf{U}^b (\mathbf{T})^T$ is proportional to the observational part of the gradient of the preconditioned cost function. We want to construct the diagonal of \mathbf{K} such that we achieve the following properties of the 4D-Var system:

1. Priority of optimisation of fluxes is ordered according to their strength.
2. Satisfactory sensitivity of the optimisation system to fluxes with small absolute amount.
3. Robustness of the system.

10 A compromise between counteracting properties 2 and 3 has to be accomplished. Identical twin experiments not presented here indicate that for constant diagonal of \mathbf{K} , the 4D-Var system EURAD-IM shows a tendency to optimise only the largest fluxes, but is too conservative for smaller fluxes and thus property 2 is not satisfied. This especially holds true for anthropogenic emissions, dominated by few very large point sources.

Using a specific construction for the standard deviation of \mathbf{K} in dependence of $\widehat{\mathbf{U}}^b := \sum_{i=1}^N \mathbf{U}_{i-1/2}^b$ (c designates a constant),

$$15 \sqrt{K_{r,r}} = c \frac{N}{\Delta t} \left[\max_s |\widehat{\mathbf{U}}^b(s)| \right]^{-\frac{1}{l}} \left[|\widehat{\mathbf{U}}^b(r)| \right]^{-\frac{l-1}{l}}, \quad (24)$$

enables an approximation of $\nabla \mathcal{J}$ which fulfills the three properties above

$$\nabla \mathcal{J}(\mathbf{v}, \mathbf{w}) \approx \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{T/2} & 0 \\ 0 & c \left(\frac{\widehat{\mathbf{U}}^b}{\max |\widehat{\mathbf{U}}^b|} \right)^{\frac{1}{l}} \end{pmatrix} \sum_{i=0}^N \begin{bmatrix} (\mathbf{T}_{0,i})^T \\ \sum_{j=1}^i (\mathbf{T}_{j-1/2,i})^T \end{bmatrix} \mathbf{H}_i^T \mathbf{R}_i^{-1} \left[\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} \right], \quad (25)$$

see Appendix B for a derivation of Eqs. (24) and (25). Equation (24) is implemented for each of the three fluxes considered, thus the constant c and $\max_s |\widehat{\mathbf{U}}^b(s)|$ are with respect to anthropogenic emissions, photosynthesis, and respiration respectively. Thus, for $l = 1$ (i.e. the diagonal of \mathbf{K} is constant) the right hand side of Eq. (25) depends linearly on \mathbf{U}^b . With increasing l the sensitivity of the system to smaller fluxes increases while the robustness decreases.

After several tests $l = 4$ has been chosen for anthropogenic emissions and $l = 2$ for photosynthesis and respiration, which is



Table 2. Configuration of background initial states and flux factors, applied for anthropogenic emissions, photosynthesis, and respiration.

configuration	background initial state	background flux factor
1	nature run + 2 ppmV	0.8
2	nature run - 2 ppmV	0.8
3	nature run + 2 ppmV	1.2
4	nature run - 2 ppmV	1.2

caused by different spatial heterogeneity of the three fluxes. The off-diagonal elements of \mathbf{K} are calculated with the diffusion approach analog to the technique outlined for matrix \mathbf{B} . A correlation of 0.4 is used between photosynthesis and respiration, whereas no correlation is assumed between anthropogenic emissions and the two biogenic fluxes.

4 Numerical experiments with synthetic data

The general concept of identical twin experiments is to execute a forward run, of which selected values are used as synthetic or artificial observations. This run is commonly called the nature run. The first guess, also called a priori run, starts thereafter with disturbed initial values and flux rates. It is also called the background run. The identical twin analysis procedure aims to assess how well the 4D-Var analysis is able to reconstruct the disturbed parameters, assimilating the given synthetic observations of varying completeness. The model run with optimised parameters is called the analysis run. Daley (1991) values the benefits of identical twin experiments, however concludes, they “err on the optimistic side”.

Numerical experiments are executed with four different background configurations listed in Table 2. For each configuration two experiments are executed. One optimising (1) only flux factors for anthropogenic emissions, biogenic respiration, and photosynthesis and another optimising (2) initial values and flux factors jointly for each grid cell. We call (1) hereinafter “only flux factor analysis” and (2) “joint analysis”.

The model domain encompasses central Europe with 5 km resolution (see Fig. 3). After a model spinup time of 30 h spanning 23 July 00 UTC to 24 July 2012 06 UTC a 12 h 4D-Var analysis run is initialised from 06-18 UTC. Synthetic measurements from the nature run are taken every 30 minutes (the model time step is 120 s) at eleven stations (see Fig. 3 and Table 1).

The observation errors are constantly 2 ppmV and the error covariance matrices \mathbf{R}_i are diagonal. The standard deviation for initial background values is $\alpha_0^b(k) \cdot c^{IV}$ (Eq. (18)). The standard deviation of the flux factors (i.e. the constant c in Eq. (24)) equals 1.4 for anthropogenic emissions, 0.6 for photosynthesis, and 0.8 for respiration respectively. The matrix \mathbf{K} is modelled with $l = 4$ (see Sect. 3.8) for anthropogenic emissions and with $l = 2$ for biogenic fluxes. The reason is that the spatial distribution of flux variations is much smoother for biogenic fluxes compared to anthropogenic emissions. While the amount of biogenic fluxes at different grid cells has mostly the same order of magnitude, the main contribution to anthropogenic emissions is caused by very few power plants, e.g. Weisweiler and Niederaußem in the Rur catchment area.



The non-diagonal entries of **B** and **K** are calculated using an influence radius of 30 km at the bottom layer, 45 km at the top of the modelled planetary boundary layer and 60 km at the top of the model domain. The L-BFGS optimisation is carried out for at most 30 iterations.

Figure 6 shows a difference plot of configuration 4 of the NmA (nature run minus joint analysis run) initial values in different vertical layers. White spots indicating compliance of the analysis and nature run. The overall initial value correction is very good with a small overestimation at the lower layers. Due to the setup of the experiment with sparse observations the meteorology is decisive for the distribution of corrections in the analysis run. Since easterly winds are predominant during the assimilation window, correction of the initial values occurs at all measurement stations and to the east of them. The vertical boundary layer rises up to 2000 m, such that correction of initial values can be seen up to this height (Fig. 6 (h)). Artificial overestimation can be seen in layer 12 (Fig. 6 (g)) southeast of Ochsenkopf. Additional tests not presented here, showed that a longer assimilation window increases this artificial deterioration of the analysis. Due to temporal high frequency measurements and the diffusive nature of atmospheric transport the LBFGS amplifies small wiggles, primarily during later iterations, resulting in an artificial over- and underestimation.

A comparison of the assimilated flux factors for configuration four, calculated once with joint analysis and once with only flux factor analysis, illustrates the benefits of the joint analysis. The results of the joint analysis in the left column of Fig. 7 are discussed first.

The flux factor analysis of anthropogenic emissions is dominated by a few but large point sources. In the Rur catchment area, due to easterly winds, the power plant Niederaußem is captured by the observation sites Jülich and Selhausen and is better reanalysed than Weisweiler, which is located a few kilometres downstream from both observation sites. The optimised flux factors of biogenic respiration (Fig. 7 (c)) and photosynthesis (Fig. 7 (e)) meet the nature run very well. Although slight underestimation persists, the flux factors of the joint analysis are mainly between background and nature run.

The analysis of the only flux factor optimisation has a stronger overestimation of anthropogenic emissions as compared to the joint analysis due to 2 ppmV higher initial values of the background run (Fig. 7 (a), (b)). Both biogenic fluxes are underestimated in large areas, as observations at early times result in a too strong forcing of photosynthesis, decreasing also biogenic respiration due to later observations. In the area of the observation stations Jülich and Selhausen, biogenic respiration is also overestimated (Fig. 7 (d)), a clear deterioration of analysis performance compared to the joint analysis.

A comparison of the cost reduction for the four configurations (listed in Table 2) for joint optimisation and only flux factor optimisation is shown in Fig. 8. For each configuration the cost function shows a stronger decrease (by a factor of ≈ 20) for the joint optimisation compared to the flux factor optimisation.

Time-series of CO₂ concentrations for configuration 4 are depicted in Fig. 9. In general a good compliance of the nature run and joint analysis at the observation sites is seen. The CO₂ concentration of the background run and the only flux factor analysis at the initial time is always two ppmV higher than the nature run. Surprisingly also the only flux factor analysis is often in accordance with the nature run, except at Selhausen, Jülich, and Hegyhatsal. As can be seen from Fig. 7 the correction of the flux factors close to Wuestebach, Cabauw, and Ochsenkopf, from the optimisation of only flux factors is wrong. This

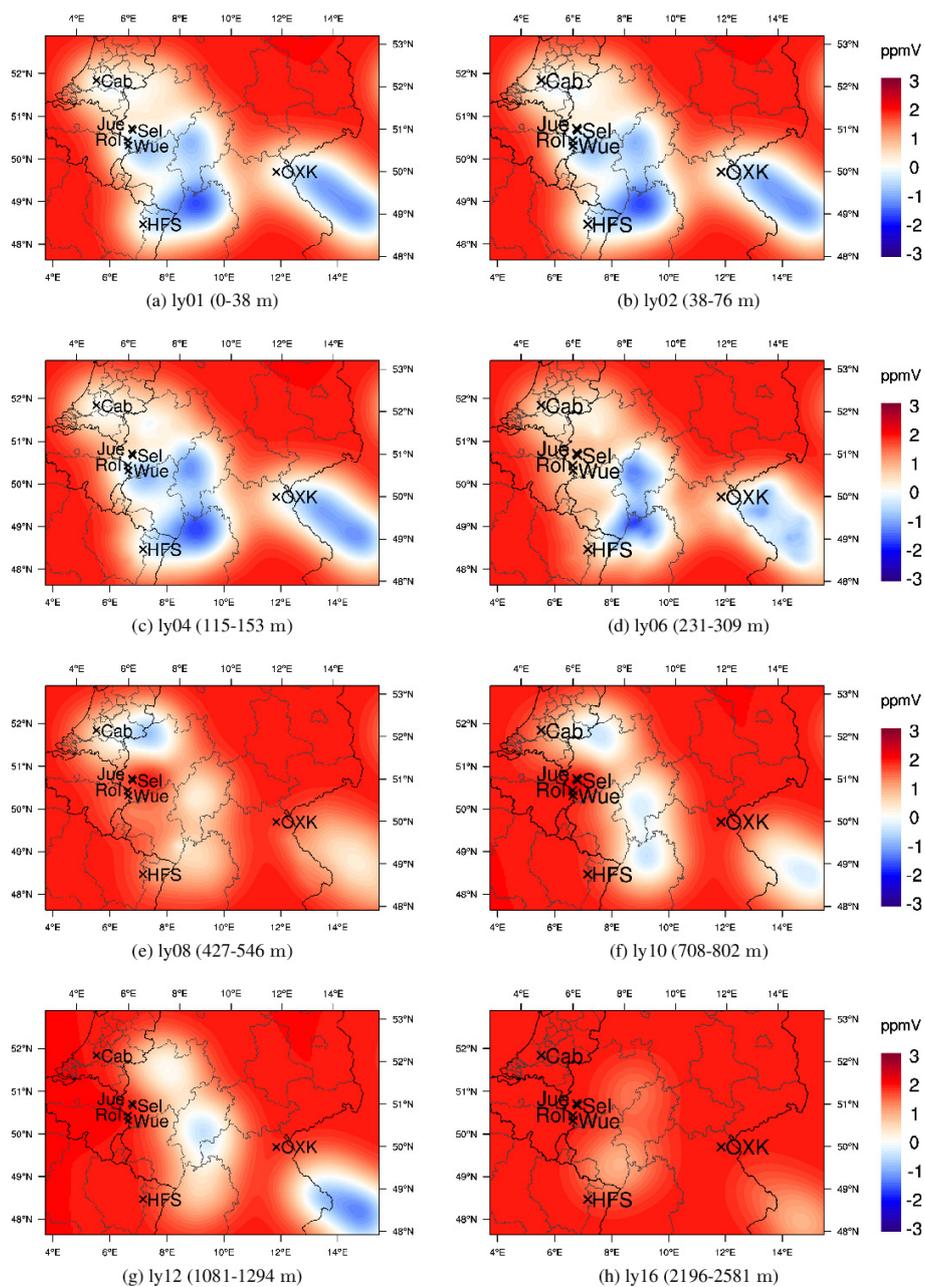


Figure 6. Zoom of the difference from initial values NmA (nature run minus joint analysis run) of configuration 4 for different layers (black crosses show synthetic observation sites). The given heights are valid for the U.S. standard atmosphere.

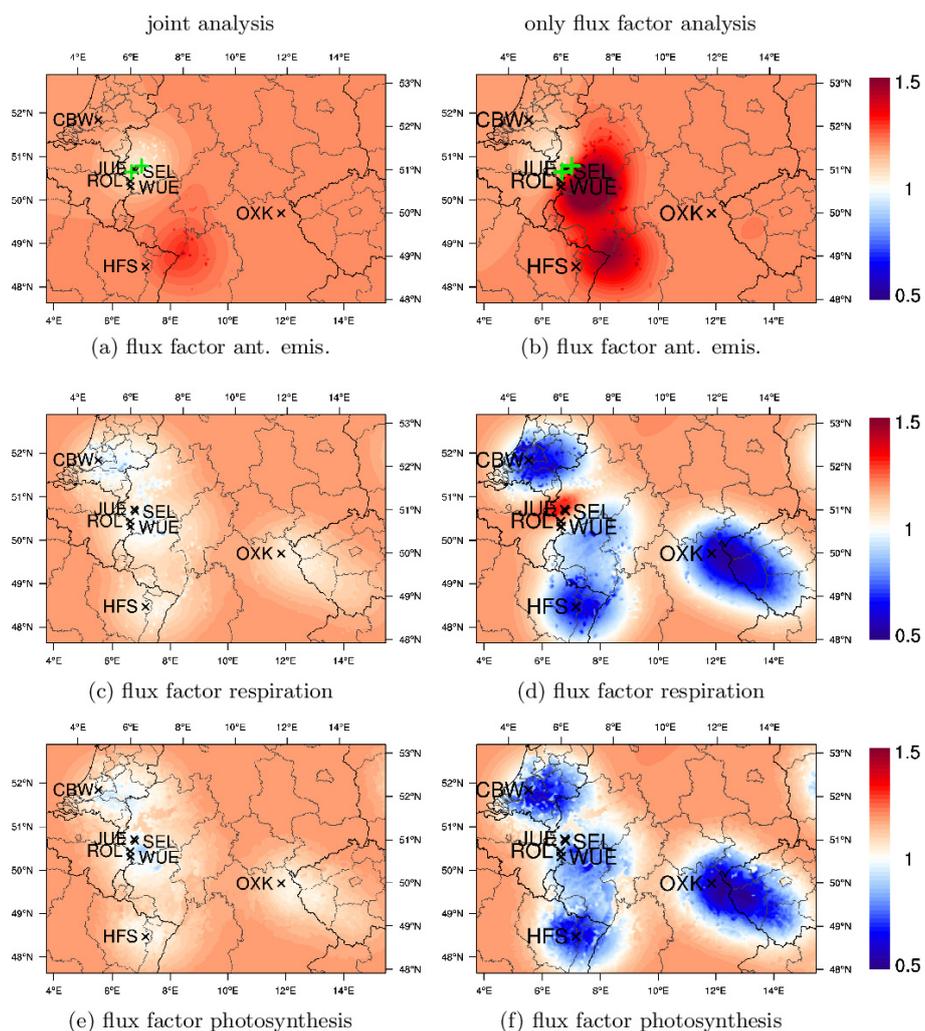


Figure 7. Analysed flux factors of anthropogenic emissions (first row), biogenic respiration (second row), and photosynthesis (third row). The analysis at the left column shows results of the joint initial value and flux factor assimilation while the right column shows an analysis which optimised only flux factors. The two biogenic fluxes are shown at the surface level, while the anthropogenic emissions are given at ≈ 270 m height, which is the layer with the highest impact due to power plants. Green plus signs indicate the location of the two biggest power plants Niederaußem and Weisweiler.

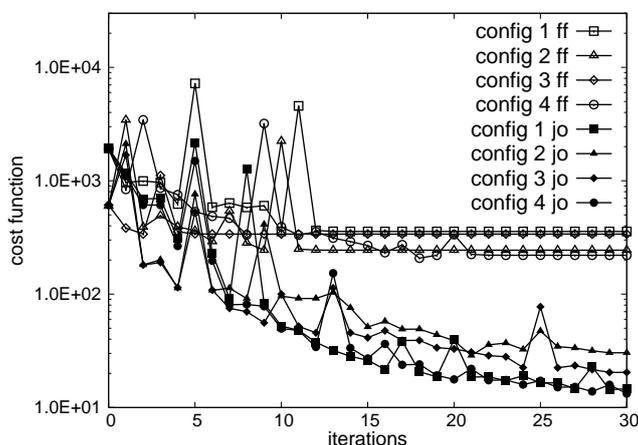


Figure 8. Calculated cost function joint (jo) and only flux factor (ff) optimisation for four configurations listed in Table 2.

shows clearly that just an evaluation of the time-series does not guarantee a sufficient analysis.

The vertical concentration profiles at the last hour of the assimilation window of the towers with measurements in higher layers is shown in Fig. 10. The vertical profiles of Cabauw, Ochsenkopf and Jülich display a slightly better concentration correction of the joint analysis compared to the only flux factor optimisation. Although the vertical profile of Hegyhatsal (Fig. 10 (c)) is very similar for the joint and the only flux factor analysis, the temporal evolution (Fig. 9 (e) and (f)) towards the last hour of the assimilation interval is quite different, indicating again that the analysed flux factors are not sufficiently well estimated.

5 Conclusions

This study assesses the potential of reducing CO_2 surface-atmosphere flux uncertainty by the proper inclusion of uncertainty of the atmospheric state with the 4D-Var method. The EURAD-IM 4D-Var model has been used to assimilate half-hourly time series of CO_2 concentration measurements by optimising initial concentration values and CO_2 fluxes jointly. CO_2 fluxes are splitted into anthropogenic emissions, biogenic respiration, and photosynthesis. These fluxes are optimised indirectly with a flux factor, changing the absolute amount of the flux, but not its temporal pattern over the assimilation window. To enable a joint optimisation of initial values and flux factors, a new calculation of the background error standard deviation for the fluxes was developed. Identical twin experiments with real meteorology and CO_2 surface fluxes are performed, using background initial values and flux factors, which are disturbed at every grid cell. A sparse observation network with eleven towers from FLUXNET is chosen, with a focus on the TR32 Rur catchment area in western Germany. The experiments are executed once with the joint optimisation of initial values and flux factors and once with the optimisation of only flux factors. Initial values are optimised up to a height of 2000 m during a 12 h assimilation run and reflect the well mixed planetary boundary layer. It is clearly shown that errors of initial values transfer to wrong estimation of flux factors. If wrong initial values are assumed to be error-free and are not part of the optimisation, the analysis downgrades. Especially biogenic fluxes, which are generally

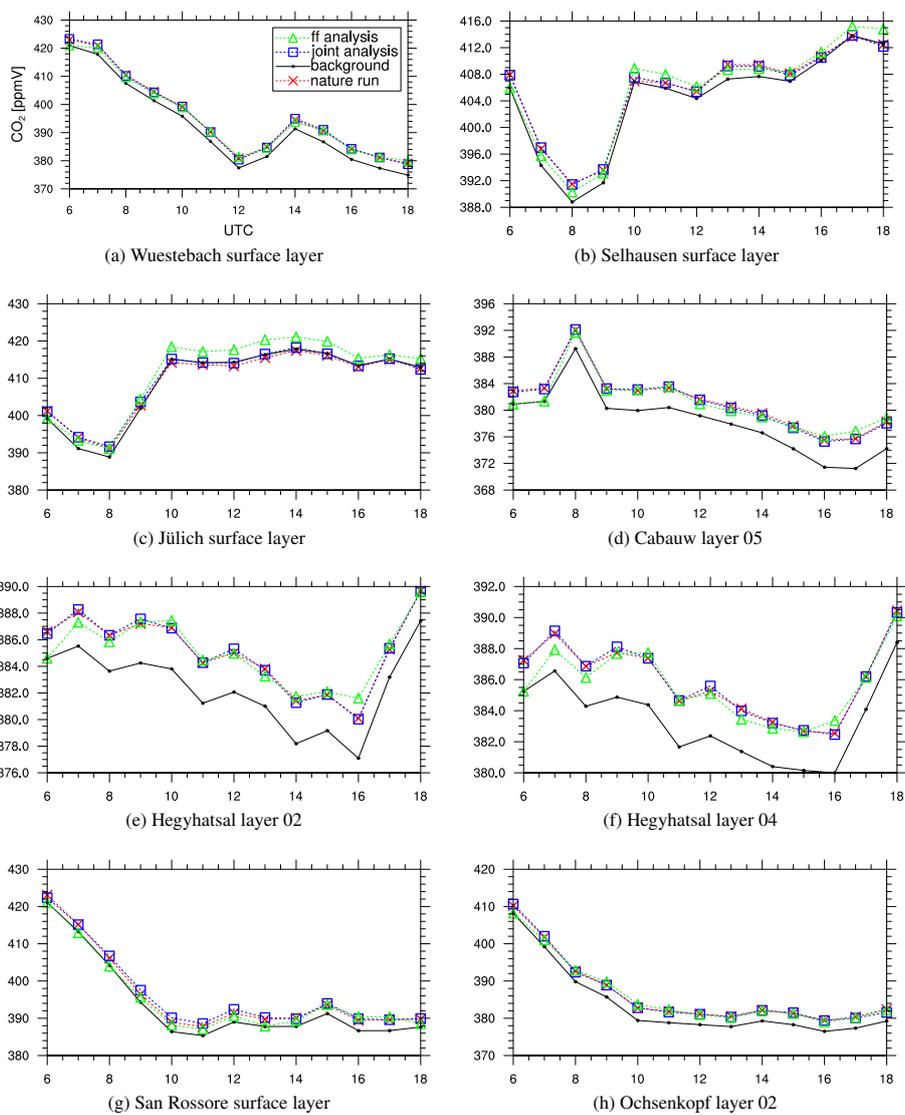


Figure 9. Time-series of CO₂ concentration of configuration four at observations sites for observations (red), background run (black), joint analysis (blue) and flux factor (FF) analysis (green).

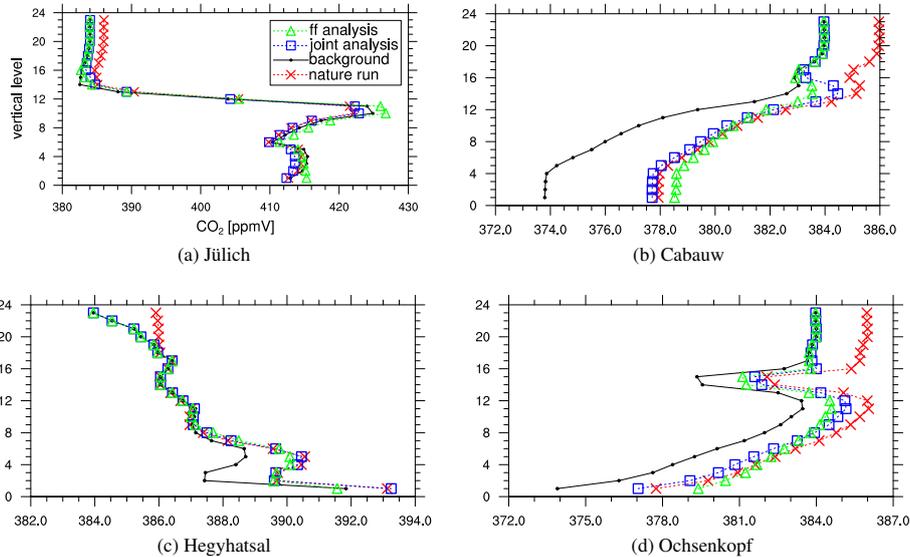


Figure 10. Vertical profile of CO₂ concentration of configuration four at observations sites with measurements also in higher layers for observations (red), background run (black), joint analysis (blue) and flux factor (FF) analysis (green) at the end of the assimilation window 24 July 2012 18 UTC.

more difficult to analyse than anthropogenic emissions, are markedly degraded. Optimising initial values and flux factors jointly improves the ability of the 4D-Var system EURAD-IM to reanalyse CO₂ fluxes.

In summary our synthetic data experiments show the impact of temporal highly resolved observations for the understanding of carbon fluxes. Wrong initial values deteriorate the ability to estimate carbon fluxes in CO₂ inversions with fine spatio-temporal resolution under the requirement, that background CO₂ fluxes capture the diurnal cycle and meteorological key parameters as planetary boundary layer and advection winds are well modelled.

15

6 Code availability

The EURAD-IM Version 5.8.1, adapted for CO₂ inversion, and modified WRFV3.6.1-CLM4.0, is developed at the Rhenish Institute for Environmental Research and stored at the Jülich Supercomputer Centre (JSC) of Research Centre Jülich and can be obtained by contacting H. Elbern (he@eurad.uni-koeln.de). Any questions should be directed to the authors.

20 Appendix A: The adjoint model (M)^T

To derive the adjoint model we combine Eq. (9) and (10) of the forward run and introduce the following notation (remember that the transport operator (see Eq. (14)) T merges the advection and diffusion operator considering the operator splitting of



the EURAD-IM: $\mathbf{T}_{i,i+1/2} = \mathbf{D}_{i,i+1/2} \mathbf{A}_{i,i+1/2}^1$, $\mathbf{T}_{i+1/2,i+1} = \mathbf{A}_{i+1/2,i+1}^2 \mathbf{D}_{i+1/2,i+1}$

$$\tilde{\mathbf{T}}_{i,i+1/2} := \begin{pmatrix} \mathbf{T}_{i,i+1/2} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}, \quad \tilde{\mathbf{T}}_{i+1/2,i+1} := \begin{pmatrix} \mathbf{T}_{i+1/2,i+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \in \mathbb{R}^{2n \times 2n}, \quad (\text{A1})$$

$$5 \quad \tilde{\mathbf{F}}_{i+1/2} := \begin{pmatrix} \mathbf{I} & \Delta t \mathbf{U}_{i+1/2}^b \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \in \mathbb{R}^{2n \times 2n} \quad (\text{A2})$$

We can now write in accordance to Eq. (9) and (10)

$$\begin{pmatrix} \mathbf{x}_{i+1} \\ \mathbf{f} \end{pmatrix} = \underbrace{\tilde{\mathbf{T}}_{i+1/2,i+1} \tilde{\mathbf{F}}_{i+1/2} \tilde{\mathbf{T}}_{i,i+1/2}}_{=\mathbf{M}_{i,i+1}} \begin{pmatrix} \mathbf{x}_i \\ \mathbf{f} \end{pmatrix}, \quad (\text{A3})$$

as $\tilde{\mathbf{F}}$ is equivalent to Eq. (10). Using Eq. (A3), the adjoint model can be derived

$$(\mathbf{M}_{i,i+1})^T \begin{pmatrix} \mathbf{x}_{i+1}^* \\ \mathbf{f}_{i+1}^* \end{pmatrix} = (\tilde{\mathbf{T}}_{i,i+1/2})^T (\tilde{\mathbf{F}}_{i+1/2})^T (\tilde{\mathbf{T}}_{i+1/2,i+1})^T \begin{pmatrix} \mathbf{x}_{i+1}^* \\ \mathbf{f}_{i+1}^* \end{pmatrix} \quad (\text{A4})$$

$$10 \quad = \begin{pmatrix} (\mathbf{T}_{i,i+1/2})^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \Delta t \mathbf{U}_{i+1/2}^b & \mathbf{I} \end{pmatrix} \begin{pmatrix} (\mathbf{T}_{i+1/2,i+1})^T & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x}_{i+1}^* \\ \mathbf{f}_{i+1}^* \end{pmatrix} \quad (\text{A5})$$

$$= \begin{pmatrix} (\mathbf{T}_{i,i+1/2})^T (\mathbf{T}_{i+1/2,i+1})^T \mathbf{x}_{i+1}^* \\ \Delta t \mathbf{U}_{i+1/2}^b (\mathbf{T}_{i+1/2,i+1})^T \mathbf{x}_{i+1}^* + \mathbf{f}_{i+1}^* \end{pmatrix} \quad (\text{A6})$$

which is Eq. (13). Equation (A6) can now be used iteratively for the calculation of $(\mathbf{M}_{0,i})^T$. For the sake of shorter notation

$$\text{we write } \mathbf{H}_i^T \mathbf{R}_i^{-1} \left[\mathbf{d}_i - \mathbf{H}_i \mathbf{M}_{0,i} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} \right] =: \begin{pmatrix} \mathbf{x}_i^* \\ \mathbf{0} \end{pmatrix},$$

$$(\mathbf{M}_{0,i})^T \begin{pmatrix} \mathbf{x}_i^* \\ \mathbf{0} \end{pmatrix} = (\mathbf{M}_{0,i-1})^T \begin{pmatrix} (\mathbf{T}_{i-1,i})^T \mathbf{x}_i^* \\ \Delta t \mathbf{U}_{i-1/2}^b (\mathbf{T}_{i-1/2,i})^T \mathbf{x}_i^* \end{pmatrix} \quad (\text{A7})$$

$$15 \quad = (\mathbf{M}_{0,i-2})^T \begin{pmatrix} (\mathbf{T}_{i-2,i})^T \mathbf{x}_i^* \\ \Delta t \mathbf{U}_{i-1-1/2}^b (\mathbf{T}_{i-1-1/2,i})^T \mathbf{x}_i^* + \Delta t \mathbf{U}_{i-1/2}^b (\mathbf{T}_{i-1/2,i})^T \mathbf{x}_i^* \end{pmatrix} \quad (\text{A8})$$

$$= \begin{pmatrix} (\mathbf{T}_{0,i})^T \mathbf{x}_i^* \\ \sum_{j=1}^i \Delta t \mathbf{U}_{j-1/2}^b (\mathbf{T}_{j-1/2,i})^T \mathbf{x}_i^* \end{pmatrix}, \quad (\text{A9})$$

which is Eq. (23).

Appendix B: Derivation of the diagonal of K

In this Sect. the specific construction of the diagonal of \mathbf{K} (Eq.(24)) and an approximation of the gradient of the cost function (Eq. (25)) is shown. First an approximation for the flux factor part of $\sum_{i=0}^N (\mathbf{M}_{0,i})^T$ (Eq. (17)) is derived with the crude



5 approximation

$$\mathbf{U}_{j-1/2}^b \approx \frac{1}{N} \sum_{i=1}^N \mathbf{U}_{i-1/2}^b =: \frac{1}{N} \widehat{\mathbf{U}}^b \quad (\text{B1})$$

Using Eq. (A9) and the notation $[\cdot]^{\text{FF}}$ as the flux factor part we can write

$$\left[\sum_{i=0}^N (\mathbf{M}_{0,i})^T \begin{pmatrix} \mathbf{x}_i^* \\ \mathbf{0} \end{pmatrix} \right]^{\text{FF}} \stackrel{(\text{A9})}{=} \Delta t \sum_{i=0}^N \left(\sum_{j=1}^i \mathbf{U}_{j-1/2}^b (\mathbf{T}_{j-1/2,i})^T \right) \mathbf{x}_i^* \quad (\text{B2})$$

$$\stackrel{(\text{B1})}{\approx} \frac{\Delta t}{N} \widehat{\mathbf{U}}^b \sum_{i=0}^N \left(\sum_{j=1}^i (\mathbf{T}_{j-1/2,i})^T \right) \mathbf{x}_i^*. \quad (\text{B3})$$

10 Using additionally the construction of the diagonal of \mathbf{K} (c designates a constant number)

$$\sqrt{K_{r,r}} = c \frac{N}{\Delta t} \left[\max_s |\widehat{\mathbf{U}}^b(s)| \right]^{-\frac{1}{t}} \left[|\widehat{\mathbf{U}}^b(r)| \right]^{-\frac{t-1}{t}}, \quad (\text{B4})$$

the gradient of the preconditioned cost function can be approximated

$$\nabla \mathcal{J}(\mathbf{v}, \mathbf{w}) \stackrel{(\text{17})}{=} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{\text{T}/2} & 0 \\ 0 & \mathbf{K}^{\text{T}/2} \end{pmatrix} \sum_{i=0}^N (\mathbf{M}_{0,i})^T \begin{pmatrix} \mathbf{x}_i^* \\ \mathbf{0} \end{pmatrix} \quad (\text{B5})$$

$$\stackrel{(\text{A9})}{=} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{\text{T}/2} & 0 \\ 0 & \mathbf{K}^{\text{T}/2} \end{pmatrix} \sum_{i=0}^N \left(\sum_{j=1}^i \Delta t \mathbf{U}_{j-1/2}^b (\mathbf{T}_{j-1/2,i})^T \mathbf{x}_i^* \right) \quad (\text{B6})$$

$$\stackrel{(\text{B3})}{\approx} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{\text{T}/2} & 0 \\ 0 & \mathbf{K}^{\text{T}/2} \left(c \frac{\Delta t}{N} |\widehat{\mathbf{U}}^b| \right) \end{pmatrix} \sum_{i=0}^N \left(\sum_{j=1}^i (\mathbf{T}_{j-1/2,i})^T \mathbf{x}_i^* \right) \quad (\text{B7})$$

$$\stackrel{(\text{B4})}{\approx} \begin{pmatrix} \mathbf{v} \\ \mathbf{w} \end{pmatrix} - \begin{pmatrix} \mathbf{B}^{\text{T}/2} & 0 \\ 0 & \left(c \frac{\widehat{\mathbf{U}}^b}{\max_s |\widehat{\mathbf{U}}^b(s)|} \right)^{\frac{1}{t}} \end{pmatrix} \sum_{i=0}^N \left(\sum_{j=1}^i (\mathbf{T}_{j-1/2,i})^T \mathbf{x}_i^* \right), \quad (\text{B8})$$

such that we get Eq. (25). As already mentioned in Sect. 3.6 and 3.8.2 the notation presented here is straightforward to apply

5 also for three fluxes but is avoided here for the sake of clarity.

Acknowledgements. We gratefully acknowledge the support by the SFB-TR32 "Pattern in Soil-Vegetation-Atmosphere-Systems: Monitoring, Modelling, and Data Assimilation" funded by the Deutsche Forschungsgemeinschaft (DFG). The authors gratefully acknowledge the computing time for the numerical experiments granted by the John von Neumann Institute for Computing (NIC) and provided on the supercomputer JURECA at Jülich Supercomputing Centre (JSC). We thank H. D. van der Gon to provide the TNO CO₂ emission inventory.

10 Further, we would like to thank M. Übel and A. Graf for support at implementing leaf respiration and soil respiration into WRF-CLM respectively. M. Memmesheimer and G. Piekorz from RIU are to be thanked for relevant discussions and technical support.



References

- Ahmadov, R., Gerbig, C., Kretschmer, R., Koerner, S., Neininger, B., Dolman, A., and Sarrat, C.: Mesoscale covariance of transport and CO₂ fluxes: Evidence from observations and simulations using the WRF-VPRM coupled atmosphere-biosphere model, *J. Geophys. Res.: Atmospheres* (1984–2012), 112, 2007.
- Ahmadov, R., Gerbig, C., Kretschmer, R., Körner, S., Rödenbeck, C., Bousquet, P., and Ramonet, M.: Comparing high resolution WRF-VPRM simulations and two global CO₂ transport models with coastal tower measurements of CO₂, *bg*, 6, 807–817, 2009.
- Baker, D., Law, R., Gurney, K., Rayner, P., Peylin, P., Denning, A., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., et al.: TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO₂ fluxes, 1988–2003, *gbc*, 20, 2006.
- Broquet, G., Chevallier, F., Rayner, P., Aulagnier, C., Pison, I., Ramonet, M., Schmidt, M., Vermeulen, A. T., and Ciais, P.: A European summertime CO₂ biogenic flux inversion at mesoscale from continuous in situ mixing ratio measurements, *J. Geophys. Res.: Atmospheres* (1984–2012), 116, 2011.
- Collatz, G., Ball, J., Grivet, C., and Berry, J. A.: Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer, *Agric. For. Meteorol.*, 54, 107 – 136, doi:[http://dx.doi.org/10.1016/0168-1923\(91\)90002-8](http://dx.doi.org/10.1016/0168-1923(91)90002-8), <http://www.sciencedirect.com/science/article/pii/0168192391900028>, 1991.
- Cooley, D., Breidt, F. J., Ogle, S. M., Schuh, A. E., and Lauvaux, T.: A constrained least-squares approach to combine bottom-up and top-down CO₂ flux estimates, *Environ. Ecol. Stat.*, 20, 129–146, 2013.
- Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, 120, 1367–1387, 1994.
- Daley, R.: *Atmospheric data analysis*, Cambridge atmospheric and space science series, Cambridge University Press, 6966, 25, 1991.
- Dolman, A., Gerbig, C., Noilhan, J., Sarrat, C., and Miglietta, F.: Detecting regional variability in sources and sinks of carbon dioxide: a synthesis, *bg*, 6, 1015–1026, 2009.
- Elbern, H. and Schmidt, H.: Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.: Atmospheres* (1984–2012), 106, 3569–3590, 2001.
- Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4D-variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Ecol. Stat.*, 15, 539 – 548, doi:[http://dx.doi.org/10.1016/S1364-8152\(00\)00049-9](http://dx.doi.org/10.1016/S1364-8152(00)00049-9), <http://www.sciencedirect.com/science/article/pii/S1364815200000499>, air pollution modelling and simulation, 2000.
- Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7, 3749–3769, doi:10.5194/acp-7-3749-2007, <http://www.atmos-chem-phys.net/7/3749/2007/>, 2007.
- Enting, I. G., Trudinger, R., and Granek, H.: Synthesis inversion of atmospheric CO₂ using the GISS tracer transport model, CSIRO Melbourne, 1993.
- Farquhar, G., von Caemmerer, S. v., and Berry, J.: A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species, *Planta*, 149, 78–90, 1980.
- Fletcher, S.: Mixed Gaussian-lognormal four-dimensional data assimilation, *Tellus A*, 62, 266–287, 2010.
- Fletcher, S. J. and Zupanski, M.: A hybrid multivariate normal and lognormal distribution for data assimilation, *Atmos. Sci. Lett.*, 7, 43–46, 2006.



- Geels, C., Gloor, M., Ciais, P., Bousquet, P., Peylin, P., Vermeulen, A., Dargaville, R., Aalto, T., Brandt, J., Christensen, J., et al.: Comparing atmospheric transport models for future regional inversions over Europe—Part 1: mapping the atmospheric CO₂ signals, *Atmos. Chem. Phys.*, 7, 3461–3479, 2007.
- 15 Gerbig, C., Körner, S., and Lin, J.: Vertical mixing in atmospheric tracer transport models: error characterization and propagation, *Atmos. Chem. Phys.*, 8, 591–602, 2008.
- Goris, N. and Elbern, H.: Singular vector based targeted observations of chemical constituents: description and first application of the EURAD-IM-SVA, *gmd Discuss.*, 8, 6267–6307, doi:10.5194/gmdd-8-6267-2015, <http://www.geosci-model-dev-discuss.net/8/6267/2015/>, 2015.
- 20 Gou, T. and Sandu, A.: Continuous versus Discrete Advection Adjoints in Chemical Data Assimilation with CMAQ, *Atmos. Env.*, 45, 4868–4881, 2011.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., et al.: Towards robust regional estimates of CO₂ sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, 2002.
- Hascoet, L. and Pascual, V.: The Tapenade Automatic Differentiation tool: principles, model, and specification, *ACM (TOMS)*, 39, 20, 2013.
- 25 Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., et al.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, *J. Geophys. Res.*: bg (2005–2012), 116, 2011.
- Kretschmer, R., Gerbig, C., Karstens, U., Biavati, G., Vermeulen, A., Vogel, F., Hammer, S., and Totsche, K.: Impact of optimized mixing heights on simulated regional atmospheric transport of CO₂, *Atmos. Chem. Phys.*, 14, 7149–7172, 2014.
- 30 Kuenen, J., Visschedijk, A., Jozwicka, M., and Denier van der Gon, H.: TNO-MACC_II emission inventory; a multi-year (2003–2009) consistent high-resolution European emission inventory for air quality modelling, *Atmos. Chem. Phys.*, 14, 10963–10976, 2014.
- Lauvaux, T., Uliasz, M., Sarrat, C., Chevallier, F., Bousquet, P., Lac, C., Davis, K. J., Ciais, P., Denning, A. S., and Rayner, P. J.: Mesoscale inversion: first results from the CERES campaign with synthetic data, *Atmos. Chem. Phys.*, 8, 3459–3471, doi:10.5194/acp-8-3459-2008, <http://www.atmos-chem-phys.net/8/3459/2008/>, 2008.
- 35 Lin, J. and Gerbig, C.: Accounting for the effect of transport errors on tracer inversions, *Geophys. Res. Lett.*, 32, 2005.
- Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, *Mathematical programming*, 45, 503–528, 1989.
- Lloyd, J. and Taylor, J.: On the temperature dependence of soil respiration, *Funct. Ecol.*, pp. 315–323, 1994.
- Memmesheimer, M., Hass, H., Tippke, J., and Ebel, A.: Modeling of episodic emission data for Europe with the EURAD emission model EEM, *Regional Photochemical Measurement and Modeling Studies*, 2, 495–499, 1995.
- 5 Nocedal, J.: Updating quasi-Newton matrices with limited storage, *Math. Comput.*, 35, 773–782, 1980.
- Oleson, K. W., Lawrence, D. M., Gordon, B., Flanner, M. G., Kluzek, E., Peter, J., Levis, S., Swenson, S. C., Thornton, E., Feddema, J., et al.: Technical description of version 4.0 of the Community Land Model (CLM), NCAR Technical Note STR, 2010.
- Peters, W., Krol, M., Van Der Werf, G., Houweling, S., Jones, C., Hughes, J., Schaefer, K., Masarie, K., Jacobson, A., Miller, J., et al.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *gcb*, 16, 1317–1337, 2010.
- 10 Peylin, P., Rayner, P., Bousquet, P., Carouge, C., Hourdin, F., Heinrich, P., Ciais, P., et al.: Daily CO₂ flux estimates over Europe from continuous atmospheric measurements: 1, inverse methodology, *Atmos. Chem. Phys.*, 5, 3173–3186, 2005.



- Pillai, D., Gerbig, C., Ahmadov, R., Rödenbeck, C., Kretschmer, R., Koch, T., Thompson, R., Neisinger, B., and Lavrié, J.: High-resolution simulations of atmospheric CO₂ over complex terrain—representing the Ochsenkopf mountain tall tower, *Atmos. Chem. Phys.*, 11, 7445–7464, 2011.
- 15 Poulitot, G., Pierce, T., van der Gon, H. D., Schaap, M., Moran, M., and Nopmongcol, U.: Comparing emission inventories and model-ready emission datasets between Europe and North America for the AQMEII project, *Atmos. Env.*, 53, 4–14, 2012.
- Rayner, P., Law, R., Allison, C., Francey, R., Trudinger, C., and Pickett-Heaps, C.: Interannual variability of the global carbon cycle (1992–2005) inferred by inversion of atmospheric CO₂ and $\delta^{13}\text{C}_{\text{CO}_2}$ measurements, *Glob. Biogeochem. Cycles*, 22, 2008.
- Rödenbeck, C., Houweling, S., Gloor, M., and Heimann, M.: Time-dependent atmospheric CO₂ inversions based on interannually varying tracer transport, *Tellus B*, 55, 488–497, 2003.
- 20 Sarrat, C., Noilhan, J., Dolman, A., Gerbig, C., Ahmadov, R., Tolk, L., Meesters, A., Hutjes, R., Ter Maat, H., Pérez-Landa, G., et al.: Atmospheric CO₂ modeling at the regional scale: an intercomparison of 5 meso-scale atmospheric models, *bg Discuss.*, 4, 1923–1952, 2007.
- Sitch, S., Friedlingstein, P., Gruber, N., Jones, S., Murray-Tortarolo, G., Ahlström, A., Doney, S. C., Graven, H., Heinze, C., Huntingford, C., et al.: Trends and drivers of regional sources and sinks of carbon dioxide over the past two decades, *bg Discuss.*, 10, 20 113–20 177, 2013.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, AVAILABLE FROM NCAR; P.O. BOX 3000; BOULDER, CO, 2008.
- Tans, P. P., Fung, I. Y., and Takahashi, T.: Observational constraints on the global atmospheric CO₂ budget, *Science*, 247, 1431–1438, 1990.
- 30 Tolk, L., Peters, W., Meesters, A., Groenendijk, M., Vermeulen, A., Steeneveld, G., and Dolman, A.: Modelling regional scale surface fluxes, meteorology and CO₂ mixing ratios for the Cabauw tower in the Netherlands, *bg*, 6, 2265–2280, 2009.
- Tolk, L., Dolman, A., Meesters, A., and Peters, W.: A comparison of different inverse carbon flux estimation approaches for application on a regional domain, *Atmos. Chem. Phys.*, 11, 10 349–10 365, 2011.
- Van Leer, B.: Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection, *J. Comp. Phys.*, 23, 276–299, 35 1977.
- Walcek, C.: Minor flux adjustment near mixing ratio extremes for simplified yet highly accurate monotonic calculation of tracer advection, *J. Geophys. Res.*, 105, 9335–9348, 2000.
- 495 Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R. A., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., et al.: Global ocean carbon uptake: magnitude, variability and trends, 2013.
- Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, *Q. J. R. Meteorol. Soc.*, 127, 1815–1846, 2001.
- Wu, X., Jacob, B., and Elbern, H.: Optimal Control and Observation Locations for Time-Varying Systems on a Finite-Time Horizon, *SIAM Journal on Control and Optimization*, 54, 291–316, 2016.
- 500 Yanenko, N. N.: The method of fractional steps, Springer, 1971.
- Zhu, Q., Zhuang, Q., Henze, D., Bowman, K., Chen, M., Yaling, L., He, Y., Matsueda, H., Machida, T., Sawa, Y., et al.: Constraining terrestrial ecosystem CO₂ fluxes by integrating models of biogeochemistry and atmospheric transport and data of surface carbon fluxes and atmospheric CO₂ concentrations, *Atmos. Chem. Phys. Discuss.*, 14, 22 587–22 638, 2014.