HAL
open science

Emission rate and chemical state estimation by 4-dimensional variational inversion

H. Elbern, A. Strunk, H. Schmidt, O. Talagrand

To cite this version:

H. Elbern, A. Strunk, H. Schmidt, O. Talagrand. Emission rate and chemical state estimation by 4-dimensional variational inversion. Atmospheric Chemistry and Physics Discussions, 2007, 7 (1), pp.1725-1783. hal-00328039

HAL Id: hal-00328039

https://hal.science/hal-00328039v1

Submitted on 18 Jun 2008

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

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

Atmos. Chem. Phys. Discuss., 7, 1725–1783, 2007
www.atmos-chem-phys-discuss.net/7/1725/2007/
© Author(s) 2007. This work is licensed
under a Creative Commons License.

Atmospheric Chemistry and Physics Discussions

Emission rate and chemical state estimation by 4-dimensional variational inversion

H. Elbern[1], A. Strunk[1], H. Schmidt[1,*], and O. Talagrand[2]

[1]Rhenish Institute for Environmental Research at the University of Cologne, Köln, Germany
[2]Laboratoire de Meteorologie Dynamique, Paris, France
[*]now at: Max-Planck-Institute for Meteorology, Hamburg, Germany

Received: 23 December 2006 – Accepted: 24 January 2007 – Published: 2 February 2007

Correspondence to: H. Elbern (he@eurad.uni-koeln.de)

Abstract

This study aims to assess the potential and limits of an advanced inversion method to estimate pollutant precursor sources mainly from observations. Ozone, sulphur dioxide, and partly nitrogen oxides observations are taken to infer source strength estimates. As methodology, the four–dimensional variational data assimilation technique has been generalised and employed to include emission rate optimisation, in addition to chemical state estimates as usual objective of data assimilation. To this end, the optimisation space of the variational assimilation system has been complemented by emission rate correction factors of 19 emitted species at each emitting grid point, involving the University of Cologne mesoscale EURAD model. For validation, predictive skills were assessed for an August 1997 ozone episode, comparing forecast performances of pure initial value optimisation, pure emission rate optimisation, and joint emission rate/initial value optimisation.

Validation procedures rest on both measurements withheld from data assimilation and prediction skill evaluation of forecasts after the inversion procedures. Results show that excellent improvements can be claimed for sulphur dioxide forecasts, after emission rate optimisation. Significant improvements can be claimed for ozone forecasts after initial value and joint emission rate/initial value optimisation of precursor constituents. The additional benefits applying joint emission rate/initial value optimisation are moderate, and very useful in typical cases, where upwind emission rate optimisation is essential. In consequence of the coarse horizontal model grid resolution of 54 km, applied in this study, comparisons indicate that the inversion improvements can rest on assimilating ozone observations only, as the inclusion of NO_x observations does not provide additional forecast skill. Emission estimates were found to be largely independent from initial guesses from emission inventories, demonstrating the potential of the 4D-var method to infer emission rate improvements. The study also points to the need for improved horizontal model resolution to more efficient use of NO_x observations.

1 Introduction

The last decade has seen increasing efforts to introduce advanced spatio–temporal data assimilation methods in atmospheric chemistry. An abundance of new earth observation data and progress in modelling skills are driving incentives to engender more reliable and comprehensive pictures of the chemical evolution of the atmosphere, mapped to regular grids. In view of the variety of information sources as given by observations, with all their heterogeneity in terms of accuracy, spatial representativity, density, frequency, and various retrieval techniques on the one hand, and model results on the other hand, advanced data assimilation and inverse modelling techniques provide the appropriate data fusion and analysis technique. Early attempts to analyse tracer fields were based on monovariate kriging techniques in the troposphere (e.g., Fedorov, 1998), and other purely spatial methods in the stratosphere (e.g., Stajner et al., 2001; Struthers et al., 2002). These methods produce chemical state estimates, frequently referred to as analyses, after assimilation of observations in model simulated fields as background.

Remote sensing earth observation data from space are mostly scattered in space and time, giving only very little information at a single time. Prerequisite for a full exploitation of these sensors is therefore some application of numerical models for spatio–temporal interpolation by assimilation of data. Attempting to combine observations of different times, intermittently applied spatial data assimilation procedures cannot make use of the known physical and chemical laws as a most useful constraint. In contrast, ability to do so, would not only enlarge the observational data base per assimilation procedure by measurements over a full time interval, but also enforce some chemical consistency, dependent on the model design.

To achieve this goal, a first successful demonstration by a stratospheric chemical box model with a small number of constituents had been provided by Fisher and Lary (1995), assessing the applicability of a variational data assimilation method. Eskes et al. (1999) applied the variational method to a two dimensional model for the assim-

ilation of total satellite columns. For the troposphere the usefulness of the variational method has been shown by Elbern et al. (1997), applying the box model version of the chemistry mechanism RADM (Regional Acid Deposition Model) (Stockwell et al., 1990). Further, the successful extension to a full chemical 4-dimensional variational (4D-var) data assimilation system could be demonstrated in the context of identical twin experimentation (Elbern and Schmidt, 1999), and for an ozone case study (Elbern and Schmidt, 2001), using the University of Cologne EURAD regional chemistry–transport model. Additional chemistry applications of the 4D–var technique were provided for both the troposphere (e.g., Chai et al., 2006) and the stratosphere (Errera and Fonteyn, 2001; Elbern et al., 2005).

In atmospheric chemistry, as is in meteorology, the parameters to be optimised by data assimilation are usually the initial state variables of the model. Hence, these initial values are implicitly assumed to be the least well known parameters and, at the same time, a critical factor for an improved analysis or forecast skill.

As chemistry transport models solve an initial–boundary value problem with strong dependencies on surface parameters, the restriction to initial value optimization is no longer justified, at least in tropospheric chemistry. For example, it is well known that, under favouring conditions, freshly emitted surface pollutants can easily enter the free and upper troposphere. In this case, better knowledge of the emission strength and meteorological stability conditions appear to be at least as important as of initial values.

A thorough assessment of uncertainties and sensitivities of ozone prediction due to uncertainties of various input parameters has been provided by various studies, e.g., Hanna et al. (1998, 2001) or Schmidt and Martin (2003). While parameters like photolysis rates and meteorological conditions are of importance, emissions still figure prominently as control parameters. At the same time, emission rates are not sufficiently well known. Especially in areas exposed to air quality problems, the errors in the emission rate estimates can be considered as among the primary causes for prediction deficiencies of pollution levels.

Independent from activities termed “data assimilation”, research on the solution of

inversion problems aiming at source and sink estimates is well established over the last decades. In most cases inversion with respect to quasi–passive tracers has been performed. Newsam and Enting (1988) and Enting and Newsam (1990) addressed the global problem of the distribution of sources and sinks of carbon dioxide by the inversion of a diffusion equation, formally solved by associated Legendre functions.

In the sequel a variety of other studies were made, all based on a very limited number of flask measurements (Bousquet et al., 1999a,b; Enting et al., 1995; Fan et al., 1998; Gloor et al., 1999; Gurney, 2002). The variational approach too has been adopted for source and sink estimates aiming to contribute to better specification of greenhouse gas budgets (Kaminski et al., 1999a,b; Houweling et al., 1999).

In order to optimise model parameters, Kaminski et al. (2002) assimilated 41 CO_2 measurement data in a simplified terrestrial biosphere model by the 4D-var technique, achieving more realistic flux simulations. Attempting to overcome the limitations of CO_2 in situ observations, satellite data from the Atmospheric Infrared Sounder (AIRS) has been assimilated into the ECMWF model using the 4D–var technique by Engelen et al. (2004). As the results were satisfying only in tropical regions, improved global source and sink estimates cannot be expected with the present data base and assimilation system configuration.

On the mesoscale, Robertson and Langner (1992) used variational data assimilation for source estimation in the frame of the ETEX experiment. By adjoint modelling, Issartel (2003) applied the concept of retroplumes for source identification and estimates within the framework of "illumination". Another approach has been selected by Bocquet (2005a,b), where the maximum entropy principle has been invoked to estimate position, time, and strengths of emission sources.

So far, all emission source studies cited above remained focused on source or sink estimates of a single passive tracer, which is observed in some way, without modelling reactive chemistry. Only few attempts have been made to address the general feasibility to solve the source inversion problem for reactive chemistry, ideally estimating precursor sources by observational data from product pollutants.

From a theoretical viewpoint, only spatio–temporal data assimilation or inversion techniques are candidates for solution, which are able to combine model information with data in a consistent way, while, at the same time, are able to provide for a Best Linear Unbiased Estimate (BLUE). Given all requested assumptions are satisfied, this property is provided by the four–dimensional variational data assimilation technique and the Kalman filter, including appropriate variants thereof.

Within the scope of an identical twin set-up, a first implementation of the 4D–variational technique for emission optimisation including reactive chemistry is described in Elbern et al. (2000). A first real world application with the EURAD model is given in Elbern and Schmidt (2002), along with a technical description. By including all emitted species at each surface grid point, the typical optimisation space of initial values by atmospheric chemical state constituents is replaced by a scaled emission rate space. A practical application on the microscale has been presented by Quélo et al. (2005) for NO_x emissions and their diurnal profile, using the Polair3D model. Adopting the variational inversion technique on the global scale, Muller and Stavrakou (2005) assimilated tropospheric column retrievals of CO and NO_2, to assess emission rates of continental scales.

Related Kalman filter implementations with sophisticated complexity reduction techniques are presented by van Loon et al. (2000), where a reduced rank square-root approach was selected to factorise covariance matrices by a few principal components (Verlaan and Heemink, 1995). Further elaboration on this technique by combination with an ensemble Kalman filter method resulted in additional skill (Hanea et al., 2004). Optimisation parameters include emission rates, photolysis rates, and deposition rates, the correction quantities of which are formally introduced as “noise” parameters in the Kalman filter formulation.

As a step toward a more comprehensive system inversion, the present study seeks to exploit the flexibility of the variational inversion technique and to combine, for the first time, emission rate and chemical state optimisation. Acknowledging the fact that area emission rates are not directly observable, inversion success can best be validated

by forecast skills, ideally by data from sites permanently withheld from assimilation. Neither for the variational approach, nor for the Kalman filter method, a systematic assessment of this kind has been provided.

Hence, it is the objective of the present paper, to

- explore the feasibility and assess the benefits of emission rate optimisation of all gaseous species emitted in a comprehensive chemistry-transport model, and
- to gain insight into the limitations, be it due to system set-up or of more fundamental nature.

As independent success criteria, assimilation analyses are validated with observations withheld from the assimilation procedure and by forecast improvements.

The paper is organised as follows: Sect. 2 exposes the underlying theory, and Sect. 3 presents the implementation of the variational assimilation system. In Sect. 4 the observation data base is given. The results of the case study are delineated in Sect. 5. Section 6 presents the conclusions.

2 Variational model inversion

The problem of finding the most probable model parameter values can be treated as a generalisation of the 4D-var approach, which is usually applied to estimate the state space variables as parameters of interest. In practical forecast applications these then serve as best known initial values (Daley, 1991; Lorenc, 1986, 1988; Talagrand and Courtier, 1987). As explained above, the focus of this paper is the inclusion of further parameters. For a review of variational parameter optimisation studies, which also consider parameters other than initial values, see for example Navon (1997). In the case of the present study, it is the emission rates which are also taken to be subject to optimisation, in addition to and joint with the initial values of the chemical constituents.

This section presents a brief outline of the theoretical background, resting on Elbern et al. (2000). The notational convention here follows as closely as possible the suggestions given in Ide et al. (1997).

We are given an a priori or first guess estimate of the chemical state vector $\boldsymbol{x}_b \in \mathbb{R}^N$, also termed background field, with N the dimension of the phase space portion for the chemical constituents, and the emission rates $\boldsymbol{e}_b \in \mathbb{R}^E$, with E the dimension of partial phase space of the emission rates. The background field is frequently obtained from a short range forecast, as in this study, or from some climatological files. The background emission rates are usually taken from emission inventories. In addition, $M(t)$ observations $\boldsymbol{y}^0(t)$ will be available at time t, scattered in a time interval $[t_0, t_T]$. The innovation vector, that is the difference between observations, available at time t, and the corresponding model equivalent state $\boldsymbol{x}_b(t)$, which evolved from the background initial value $\boldsymbol{x}_b(t_0)$ and emission rates $\boldsymbol{e}_b(t)$, is denoted

$$\boldsymbol{d}(t) = \boldsymbol{y}^0(t) - \mathcal{H}(t)\boldsymbol{x}_b(t). \tag{1}$$

The forward observation operator $\mathcal{H}$ maps from model space to observation space, producing the model equivalents of oberservations, given at the time t.

The innovation vector $\boldsymbol{d}$, the deviations from the background chemical state $\delta\boldsymbol{x}(t_0) := \boldsymbol{x}(t_0) - \boldsymbol{x}_b(t_0)$, and a suitably defined perturbation function of emissions $\delta\boldsymbol{u}=u(\boldsymbol{e}, \boldsymbol{e}_b)$, scaling the deviation of modified emissions $\boldsymbol{e}$ from background emission rate values $\boldsymbol{e}_b$, are combined in a quadratic form to define an incremental formulation of a cost function, objective function or distance function $\mathcal{J}$ as follows (Courtier et al., 1994):

$$\mathcal{J}(\delta\boldsymbol{x}(t_0), \delta\boldsymbol{u}) = \frac{1}{2}(\delta\boldsymbol{x}(t_0))^T\mathbf{B}^{-1}\delta\boldsymbol{x}(t_0) + \frac{1}{2}(\delta\boldsymbol{u})^T\mathbf{K}^{-1}\delta\boldsymbol{u} + \frac{1}{2}\sum_{i=0}^{N}(\boldsymbol{d}(t_i) - \mathbf{H}(t)\delta\boldsymbol{x}(t_i))^T\mathbf{R}^{-1}(\boldsymbol{d}(t_i) - \mathbf{H}(t_i)\delta\boldsymbol{x}(t_i)), \tag{2}$$

$\mathcal{J}$ is a scalar functional defined on the time interval $t_0 \le t \le t_N$ dependent on the vector

valued state variable $\boldsymbol{x}(t)$. $\mathbf{H}(t) \in \mathbb{R}^{M(t)\times N}$ is a linearised approximation of the forward observation operator $\mathcal{H}$. The error covariance matrix of the first guess or background values $\boldsymbol{x}_b$ is denoted $\mathbf{B} \in \mathbb{R}^{N\times N}$, while error covariance matrices of emission perturbation functions $\delta\boldsymbol{u}$ and of observations $\boldsymbol{y}(t)$ are denoted $\mathbf{K} \in \mathbb{R}^{E\times E}$ and $\mathbf{R} \in \mathbb{R}^{M(t)\times M(t)}$, respectively.

Chemical tendencies as prescribed by a chemistry transport model (CTM) and by emissions, are given by

$$\frac{d\boldsymbol{x}}{dt} = \mathcal{M}(\boldsymbol{x}) + \boldsymbol{e}(t), \tag{3}$$

where $\mathcal{M}$ acts as a generally nonlinear model operator. Both $\boldsymbol{x}(t_0)$ and $\boldsymbol{e}(\tau), \tau \in [t_0, t]$ control the state variable $\boldsymbol{x}(t)$ at time t. For a chemistry-transport model as the EURAD model applied in this study, the differential equation can be written as (Elbern and Schmidt, 2001):

$$\frac{\partial c_i}{\partial t} + \nabla\cdot(\boldsymbol{v}c_i) - \nabla\cdot(\rho\mathbf{G}\nabla\frac{c_i}{\rho}) - \sum_{r=1}^{R}\left(k(r)\left(s_i(r_+) - s_i(r_-)\right)\prod_{j=1}^{U} c_j^{s_j(r_-)}\right) = E_i + D_i \tag{4}$$

where c_i is the concentration of species i, $\boldsymbol{v}$ is the wind velocity, $s \in \mathrm{N}_0$ is the stoichiometric coefficient, $k(r)$ is the reaction rate of reaction r, either being productive (r_+) or destructive (r_-) for species i, U is the number of species in the mechanism, E_i is the emission rate of species i, D_i deposition rate of species i, the air density is denoted by ρ, and $\mathbf{G}$ is the symmetric eddy diffusivity tensor.

Now let $\mathbf{M}'$ be the tangent linear model operator of $\mathcal{M}$. The evolution of perturbation $\delta\boldsymbol{x}$ from $\boldsymbol{x}_b(t)$ and $\boldsymbol{e}(t)$ follows from the tangent linear form of (3)

$$\frac{d\delta\boldsymbol{x}}{dt} = \mathbf{M}'\delta\boldsymbol{x}(t) + \delta\boldsymbol{e}(t). \tag{5}$$

Aiming to optimise initial values $\boldsymbol{x}_b(t_0)$ and emission rates $\delta\boldsymbol{e}(t)$ jointly, both parameters must be combined in a common vector by suitable scaling. This will be accomplished

by introducing a function $\boldsymbol{u} = \boldsymbol{u}(e)$. The relation between perturbation $\delta\boldsymbol{u}$ and the perturbation of emission rates $\delta\boldsymbol{e}(t)$ is implementation dependent and will be specified later. The composit vector $\delta z := (\delta\boldsymbol{x}^T(t_0), \delta\boldsymbol{u}^T)^T$ is the full control parameter of the model evolution.

For the related integration operator or resolvent, performing a model integration from time t_0 to time t, we can reformulate

$$\delta\boldsymbol{x}(t) = \tilde{\mathbf{M}}(t, t_0)\,\delta z, \tag{6}$$

where $\tilde{\mathbf{M}}(t, t_0)$ is the integration operator from time t_0 to t. In order to minimise $\mathcal{J}$ by gradient descent or quasi–Newton methods efficiently, we want to determine the gradient of $\mathcal{J}$ with respect to the joint chemical state and emission rate variable δz, that is $\partial\mathcal{J}/\partial\delta z$. The gradient of the cost function $\mathcal{J}$ then reads

$$\partial\mathcal{J}/\partial z = \mathbf{B}^{-1}\delta\boldsymbol{x}(t_0) + \mathbf{K}^{-1}\delta\boldsymbol{u} - \sum_{t_i=t_0}^{t_N} \mathbf{H}^T(t_i)\mathbf{M}^T(t_0, t_i)\mathbf{R}^{-1}\left(\boldsymbol{d}(t_i) - \mathbf{H}(t_i)\delta\boldsymbol{x}(t_i)\right). \tag{7}$$

Here, $\mathbf{H}^T$ and $\mathbf{M}^T$ denote the adjoint of the tangent–linear observation operator $\mathbf{H}$ and model $\mathbf{M}'$. With the costs $\mathcal{J}$, the perturbation fields δz, and $\partial\mathcal{J}/\partial\delta z$ once calculated, the minimisation routine can be processed, resulting in a further step toward a better estimate $\delta z_a := (\delta\boldsymbol{x}_a^T(t_0), \delta\boldsymbol{u}_a^T)^T$, expected to converge to the best linear unbiased chemical state of the atmosphere and emission rate, provided that the tangent linear approximation is sufficiently valid.

The adjoint formulation of (4) then reads, after application of the variational calculus

$$-\frac{\partial\delta c_i^*}{\partial t} - \boldsymbol{v}\nabla\delta c_i^* - \frac{1}{\rho}\nabla\cdot(\rho\mathbf{K}\nabla\delta c_i^*) + \sum_{r=1}^{R}\left(k(r)\frac{s_i(r_-)}{c_i}\prod_{j=1}^{U} c_j^{\,s_j(r_-)}\sum_{n=1}^{U}\left(s_n(r_+) - s_n(r_-)\right)\delta c_n^*\right) = 0 \tag{8}$$

with δc_i^* being the adjoint variable of c_i, while D_i is held fixed.

ACPD
7, 1725–1783, 2007

EGU

3 Implementation of the EURAD 4D–var system

The description of the components of the EURAD 4D–var system follows the algorithms exposed in the previous section. It includes (i) the EURAD model $\mathcal{M}$ and its adjoint $\mathbf{M}^T$, (ii) the formulation of both background error covariance matrices **B** and **K** for the initial states and the emission factors, respectively, and their treatment to precondition the minimisation problem, (iii) the observational basis and its related error covariance matrix, and (iv) the minimisation including the transformation for preconditioning.

3.1 The EURAD forward model CTM2 and its adjoint

The chemistry transport model from which the adjoint version is developed is the University of Cologne EURopean Air pollution Dispersion model (EURAD) (Hass et al., 1995; Elbern and Schmidt, 2001), which is an early offspring of the Regional Acid Deposition Model RADM2 (Chang et al., 1987).

The chemistry transport model calculates the transport, diffusion, and gas phase transformation of 60 chemical species with 158 reactions. These processes are calculated sequentially by a symmetric operator splitting technique, when stepping from t to $t+\Delta t$ (Yanenko, 1971; McRae et al., 1982). This approach is shown to minimise systematic biases as introduced by a fixed sequence in the splitting technique. In the present configuration the following operator sequence is implemented:

$$\boldsymbol{x}^{t+\Delta t} = T_h T_z D_z A D_z T_z T_h \boldsymbol{x}^t, \tag{9}$$

where T, D denote transport and diffusion operators in horizontal (h) or vertical (z) direction, respectively. The parameterisation of the emission sources and deposition processes are included in the gas phase chemistry module A and vertical diffusion D_z, respectively. The dynamic time step Δt of the advection operators is 10 min. The Bott (1989) upstream algorithm is chosen to calculate the horizontal and vertical advection. The vertical diffusion is semi-implicitly discretised following Crank–Nicholson, with the Thomas algorithm used as solver.

In the present configuration a semi-implicit and quasi steady state approximation method (QSSA) is applied for the numerical solution of the gas phase chemistry stiff ordinary differential equation system, following Hesstvedt et al. (1978). The chemistry time step Δt_c of the stiff ordinary differential equation solver is highly variable in time, the calculation of which follows McRae et al. (1982), with a limited set of species selected to determine the time step. By practical reasons a lower bound is defined at 1/50 min, while the upper bound is given by the dynamic time step of 10 min.

In this study the radiative transfer equation solver of Madronich (1987) is applied as preprocessor to the CTM and its adjoint.

The integration domain applies the Lambert conformal projection centered at 50° N latitude, 10° E longitude. A horizontal resolution of 54 km with 77 grid points in the x-direction and 67 grid points in the y-direction is employed. The model's horizontal grid structure is defined by the "Arakawa C" grid stencil. In the vertical 15 levels with terrain–following σ coordinates of Lorenz type are used, with refinements at the lowest levels. The lowest model half layer, where concentrations, temperature and winds are given, is chosen to represent 38 m height. The isobaric level of 100 hPa which defines the top of the model, is taken as a material surface. Grid size and number of species result in a state space dimension $\dim(\delta \boldsymbol{x}) \approx 4.6 \times 10^6$.

The initial model state, that is, the initial values prior to the spin-up run at the onset of a case study, includes a seasonal mean concentration of longer lived species dependent on latitude and height for the first forward model run, covering 1–2 August, while serving as model spin-up period. Inflow boundary values are defined in the same way as the initial model state. All later model runs start with the simulated final model state of the preceeding run.

The emission module includes 19 emitted species, formally available at each emitting grid-point. Emissions into other layers than the lowest are simulating injections due to stack overshooting (Briggs, 1975). Emission rates are distributed over typical diurnal cycles for working days, saturdays, and sundays. Predefined diurnal cycles are taken as a priory knowledge, shaping the emission rates over the day. Examples of working

day diurnal cycles of six species are given in Fig. 1. With this configuration, a phase space dimension of the emission rates of $\dim(\delta \boldsymbol{e}) = \mathcal{O}(10^5)$ results.

The emission data in this study are taken from EMEP (co-operative programme for monitoring and evaluation of the long range transmission of air pollutants in Europe) and further processed as presented in Memmesheimer et al. (1995). Processing includes also the seasonal and diurnal redistribution, as well as attributions to working days, saturdays, and sundays. Therefore, emission correction factors inferred in this study by a 3 week case study cannot serve as any validation of the EMEP inventory for the year 1997. The deposition modelling follows the method proposed by Wesely (1989).

The adjoint of the EURAD model can be developped from the adjoint differential Eq. (8), from the adjoint of the numerical solvers of the forward model (4), or from the forward code. In this study the latter approach is adopted, which comprises the coding and implementation of the adjoint operators of T_h, T_z, A and D_z given in (9). The adjoint chemistry was coded by hand, while for the advection and diffusion routines the AMC adjoint model compiler (Giering and Kaminski, 1998) and O∂yssée differentiation system (Rostaing et al., 1993; Faure and Papegay, 1998) has been used. The correctness of the adjoint code was checked by the method proposed by Chao and Chang (1992). For the gas phase chemistry solver and the implicit vertical diffusion operator, which apply adaptive time step techniques, the same time steps are taken for backward integration as determined by the forward integration.

As meteorological driver the Penn State/NCAR mesoscale model MM5 is applied. For a comprehensive description of the model see Anthes et al. (1987) for the precursor version MM4 and for an extended update description see Grell et al. (1993). MM5 is a primitive equation model, used in this study with an integration domain encompassing the area from the Mediterranean Sea to North Norway and from the East Atlantic to European Russia. For the presented simulations the hydrostatic mode of MM5 is taken. The horizontally staggered grid for surface pressure, temperature, water vapour, and horizontal velocity is based on the Arakawa B-grid scheme with a resolu-

tion of 54 km. Key parameterisations for the simulation of air mass transport include Blackadar mixing-layer parameterisations, Kuo cumulus parameterisation, fourth order horizontal diffusion, and K-theory vertical diffusion with Richardson number dependent-coefficient G_{zz} above the mixed layer.

A detailed application of MM5 as part of the EURAD model configuration is given in Jakobs et al. (1995).

3.2 Preconditioning of the cost function

According to (2), the present implementation requires information about the background error covariances of both the system state variables (= "tracer concentrations") and the emission rates. While it is obvious that there are correlations between concentration levels of emitted species and their emission rates in the vicinity of sources, as well as further correlations after the action of chemical transformation and transport processes, the implementation does not yet include cross-covariances between the initial state and the emission rates. This is evident from the formulation of (2). This is due to the facts that even the numerical treatment of **B** is computationally challenging, given a space state dimension of 4.6×10^6, that the cross–correlations are highly dependent on weather conditions, and that sufficient information on this can hardly be compiled within a limited case study.

In Elbern and Schmidt (2001) a first implementation of chemical 4D-variational data assimilation by singular value decomposition was presented, where **B** has been furnished with off-diagonal elements for spatial correlations. It is also pointed out, that the condition number of **B** becomes sensibly unfavourable for efficient minimisation with extending radius of influence and smoothness of the structure functions. As a consequence, preconditioning the minimisation problem becomes increasingly difficult. An effective procedure to introduce preconditioning is by transformation of the optimisation parameters by square roots of **B** and **K**. With square root factoriszations $\mathbf{B}=\mathbf{B}^{1/2}\mathbf{B}^{T/2}$

and $\mathbf{K}=\mathbf{K}^{1/2}\mathbf{K}^{T/2}$ following Courtier (1997), we define new variables $\boldsymbol{v}$ and $\boldsymbol{w}$ by

$$\boldsymbol{v} := \mathbf{B}^{-1/2}\delta\boldsymbol{x}, \qquad \boldsymbol{w} := \mathbf{K}^{-1/2}\delta\boldsymbol{u}, \tag{10}$$

leading to a minimisation problem equivalent to Eq. (2).

The cost function then reads

$$J(\boldsymbol{v},\boldsymbol{w}) = \frac{1}{2}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{v} + \frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w} + \frac{1}{2}\sum_{t_i=t_0}^{t_N}\left(\boldsymbol{d}(t_i) - \mathbf{H}\delta\boldsymbol{x}(t_i)\right)^{\mathrm{T}}\ \mathbf{R}^{-1}\ \left(\boldsymbol{d}(t_i) - \mathbf{H}\delta\boldsymbol{x}(t_i)\right). \tag{11}$$

The gradient of J with respect to $(\boldsymbol{v},\boldsymbol{w})^T$ can be shown to be

$$\nabla_{(\boldsymbol{v},\boldsymbol{w})^T} J = \begin{pmatrix}\boldsymbol{v}\\ \boldsymbol{w}\end{pmatrix} - \begin{pmatrix}\mathbf{B}^{1/2} & \mathbf{0}\\ \mathbf{0} & \mathbf{K}^{1/2}\end{pmatrix} \times \sum_{i=0}^{N}\mathbf{M}^T(t_0,t_i)\mathbf{H}^T\mathbf{R}^{-1}(\boldsymbol{d}(t_i) - \mathbf{H}\delta x(t_i)), \tag{12}$$

The transformation (10) efficiently compensates for the specific part of ill-conditioning introduced by any formulation of **B** and **K**.

3.3 Background error covariance matrix **B**

Due to its size, formally comprising $\mathcal{O}(10^{12})$ entries, the background error covariance matrix **B** has to be limited to a few principal components only, like leading singular vectors (e.g., Elbern and Schmidt, 2001). The generalisation of **B** to anisotropic and inhomogeneous radii of influence by an explicit covariance matrix model without relaxing the required property of positive definiteness is not straightforward (Hölzemann et al., 2001). In that paper an explicit inversion in observation space is chosen, which

is hardly practical in three dimensions. In addition, in the context of variational data assimilation, the preconditioning of the minimisation procedure requires the square root of the covariance matrix. For optimisation problems with a dimension as high as the problem at hand, both requirements can only be met by choosing a proper covariance model operator, rather than a full matrix. Hence, in contrast to the prior studies, and following Weaver and Courtier (2001) with the promise of higher flexibility in designing anisotropic and hetereogeneous influence radii, the latter option is implemented in this work. We need to define $\mathbf{B}$ as an operator, which can be easily factorised as $\mathbf{B}^{1/2}\mathbf{B}^{T/2}$, to account for the preconditioning requirements. Weaver and Courtier (2001) show that the diffusion equation serves as a valid operator for square-root covariance operator modelling, with flexibility to account for inhomogeneous and anisotropic correlation length, by suitable adjustments of local diffusion coefficients. Additionally, the diffusion equation is self-adjoint. Therefore, the operator can be easily split into $\mathbf{B}^{1/2}\mathbf{B}^{T/2}$ by applying only half the integration time of the diffusion equation.

The following operator splitting scheme is implemented as covariance filter, the single elements of which will be described in detail later:

$$\mathbf{B} = \mathbf{\Sigma}\,\mathbf{C}\,\mathbf{\Sigma} \tag{13}$$

$$\mathbf{C}^{1/2} = \mathbf{\Lambda}\mathbf{L}_v^{1/2}\mathbf{L}_h^{1/2}\mathbf{W}^{-1/2} \tag{14}$$

$$\delta\boldsymbol{x} = \mathbf{\Sigma}\mathbf{\Lambda}\mathbf{L}_v^{1/2}\mathbf{L}_h^{1/2}\mathbf{W}^{-1/2}\boldsymbol{v} \tag{15}$$

$\mathbf{\Sigma}$ is the diagonal matrix of background-error standard deviations, $\mathbf{C}$ is the covariance filter, with $\mathbf{\Lambda}$ a normalisation operator and $\mathbf{L}_{h,v}$ the horizontal and vertical diffusion operators, respectively. The diagonal matrix $\mathbf{W}$ includes correction factors needed for the grid, which accounts for the changing heights of the grid cells due to the application of σ-coordinates.

3.3.1 Background error standard deviations

In 4D-var, the background errors of the chemical species, that is the standard deviations $\sqrt{\mathrm{diag}(\mathbf{B})}$, must be selected a priori, but are amenable to posteriori by χ^2–validation following Talagrand (1998). After several test runs, the standard deviations are assumed to be dependent of species and height, the former accounting for the different degree of variability of the individual constituents, and the latter reflecting the decreasing confidence in knowledge on the chemical states with height. Introducing a height level k dependent relative error $\epsilon_{\mathrm{rel}}(k)$ and a species l dependent absolute error $\epsilon_{\mathrm{abs}}(l)$, the following formulation is selected, with the double indices indicating height and species dependence for the square roots of diagonal elements. Here

$$\sqrt{(B_{l,l,k,k})} = \max\left(x_{\mathrm{b}}(l,k)\cdot\epsilon_{\mathrm{rel}}(k), \epsilon_{\mathrm{abs}}(l)\right) \tag{16}$$

where the relative error is modelled as

$$\epsilon_{\mathrm{rel}}(k) = 1/2\cdot\exp\left(0.69\frac{1-k}{1-\mathrm{kmax}}\right), \tag{17}$$

with kmax being the number of model levels. Equation 17 implies that the relative background error is increasing exponentially with model height from 50% to 100%. The absolute background errors $\epsilon_{\mathrm{abs}}(l)$ for the different species taken are given in Table 1.

3.3.2 Radii of influence

While the diffusion approach for covariance modelling allows for highly anisotropic and inhomogeneous radii of influence or (de-)correlation length L, the related information requested is not available in 4D-var, and must be estimated in some way. In Elbern and Schmidt (2001), assimilation performance dependencies have been investigated for different influence radii of up to more than 150 km, but the result was found to be of moderate sensitivity only. This is due to the fact that the adjoint backward integration over a longer time span gains properties to disperse information spatially as well, which

easily overrides formally introduced small radii of influence, that is, a few grid cell units. This situation is given, where radii no larger than 3 grid cells (L=162 km) are selected.

The horizontal radii of influence are defined to be increasing with height. Figure 2 shows height-dependent influence radii for calculation of the corresponding diffusion parameters. The procedure to infer the diffusion coefficients $\kappa_h(i,j)$ and $\kappa_v(k)$ and the corresponding number of time steps has to comply with the following conditions: As demonstrated in Weaver and Courtier (2001), L is related to the pseudo-diffusion coefficient κ and integration time T by

$$L = \sqrt{2\kappa T}. \tag{18}$$

On the other hand, the stability condition for explicit solvers of diffusion equations as parabolic differential equations, requires time stepping constrained by

$$\Delta t_h \leq \frac{(\Delta x)^2}{2\ \max_{i,j}(\kappa_h(i,j))}, \qquad \Delta t_v \leq \frac{\min_{\Delta\sigma(k)}(\Delta\sigma(k))^2}{2\ \max_k(\kappa_v(k))} \tag{19}$$

for horizontal and vertical pseudo-diffusion, respectively. Δx and $\Delta\sigma(k)$ denote horizontal and level k dependent vertical grid spacing. For the uppermost level, a horizontal correlation length of L_{top}=250 km is estimated, while for the surface layer L_{bot}=54 km are taken. The vertical interpolation rule is chosen to be linear on the bi-sectioned σ-level height scale, taking the planetary boundary height σ_{pbl} as inner limit and the corresponding correlation length is L_{pbl}=80 km.

In contrast to the horizontal correlation length, the vertical variability of the vertical correlation length L_v is inferred from the vertical diffusion coefficient κ_v, as provided by the preprocessed MM5 run. This ensures features like well mixed boundary layers and related heights being well reflected by vertical correlation lengths. Therefore, a dynamical control of the vertical correlation length is provided at no additional computational costs.

3.3.3 Normalisation matrix

The normalisation operator $\mathbf{\Lambda}$ is a diagonal matrix counteracting the flattening due to the action of the diffusion operators $\mathbf{L}$. It is ensured that the background errors on the main diagonal of $\mathbf{B}$ are those given by $\mathbf{\Sigma}$.

Following Weaver and Courtier (2001), the normalisation matrix has to be calculated for each grid cell at analysis time t_0, since vertical diffusion depends on varying meteorological conditions. The present application accomplishes this by two options:

- Applying the diffusion filter consisting of horizontal and vertical diffusion to input vectors $\boldsymbol{e}_l=(0,\ldots,0,1,0,\ldots,0)^{\mathrm{T}}$, the entry equal to one being located at grid point l. For each grid point l the normalisation factor then reads:

$$(\mathbf{\Lambda})_{ll} = \frac{1}{\sqrt{t_l}} \qquad \text{with} \tag{20}$$

$$t_l = \boldsymbol{e}_l^{\mathrm{T}}\mathbf{L}_{\mathbf{v}}^{1/2}\mathbf{L}_{\mathbf{h}}^{1/2}\mathbf{W}^{-1}\mathbf{L}_{\mathbf{h}}^{\mathrm{T}/2}\mathbf{L}_{\mathbf{v}}^{\mathrm{T}/2}\boldsymbol{e}_l. \tag{21}$$

- Alternatively, a randomization method Fisher and Lary (1995) can be applied, where a set of Q Gaussian random vectors $\boldsymbol{v}_q \in \mathbb{R}^N$ are generated, with the statistical expectations $\mathcal{E}(\boldsymbol{v}_q)=\mathbf{0}$ and $\mathcal{E}(\boldsymbol{v}_q\boldsymbol{v}_q^T)=\mathbf{I}, \quad q=1,\ldots,Q$. With the square root of the covariance filter $\boldsymbol{v}_q=\mathbf{L}_v^{1/2}\mathbf{L}_h^{1/2}\mathbf{W}^{-1/2}\boldsymbol{v}_q$ we then obtain

$$t_i^{-2} \approx \mathrm{diag}_i\left(\frac{1}{Q-1}\sum_{q=1}^{Q}\boldsymbol{v}_q\boldsymbol{v}_q^T\right). \tag{22}$$

The estimated randomization error is then $=1/\sqrt{2Q}$.

3.3.4 Emission parametrisation and covariance matrix

Formally, emission rates can vary for each species l at each emitting grid cell (i, j, k) and at each time step t_i. Allowing for stepwise emission variations results in an extremely ill-posed inversion problem. Therefore, the degree of freedom of the emission rate space state can be drastically reduced by taking the diurnal profile shape as strong constraint, as this is comparatively well known. By this adherence to diurnal cycle shapes, given a priori, only the amplitudes can be taken as control parameter subject to optimisation. Further, positive definiteness of emissions must be enforced. Presevation of both, diurnal cycle shape and positive definiteness define the function $\delta \boldsymbol{u} = u(\boldsymbol{e}, \boldsymbol{e}_b)$ as

$$\delta \boldsymbol{u} := \ln(\boldsymbol{e}) - \ln(\boldsymbol{e}_b), \tag{23}$$

where the logarithms are taken componentwise at each location and for each emitted species, that is $\delta u(i, j, k, l) = \ln(e(i, j, k, l)/e_b(i, j, k, l)) =$ const. throughout the assimilation interval.

We set $\delta \boldsymbol{e}(t) = : \boldsymbol{e}(t) - \boldsymbol{e}_b(t)$. For notational convenience, grid and species indices are temporarily mapped on vector index $(i, j, k, l) \rightarrow s$. Then, with the diagonal matrix $\mathrm{diag}(\mathbf{U})_s := \exp(\delta u_s) - 1$, Eq. (5) can be reformulated in terms of the emission parametrisation

$$\frac{d\delta \boldsymbol{x}}{dt} = \mathbf{M}' \delta \boldsymbol{x}(t) + \mathbf{U} \boldsymbol{e}_b(t). \tag{24}$$

Hence, optimisation of $\delta \boldsymbol{u}$ results in a correction factor $f(i, j, k, l) = e(i, j, k, l)/e_b(i, j, k, l)$, which is location and species dependent, yet time independent.

In the progression of the case study, the emission inventory values are taken as background $\boldsymbol{e}_b$ only for the first assimilated day. For later days background values $\boldsymbol{e}_b$ are taken from the optimized emission correction factors of the preceeding assimilation cycle.

The background emission rate covariance matrix $\mathbf{K}$ is specified as block diagonal matrix by sub-matrices $\tilde{\mathbf{K}}(i, i, j, j)$ for each surface grid cell (i, j), describing emission rate covariances between emitted species at each location. The construction is implemented by first estimating $\tilde{\mathbf{K}}^{1/2}$. The information used is a table of the annual emitted amount of NO_x, SO_x and volatile organic compounds (VOC), splitted into different polluter groups as provided by the EMEP inventory: http://www.emep.int. Emission error correlations have been coded by squaring $\tilde{\mathbf{K}}^{1/2}$ and composing the full $\mathbf{K}$, granting positive definiteness of $\mathbf{K}$ and enabling transformation (10). The upper triangle sub-matrix entries are presented in Fig. 3.

Background emission standard deviations are assumed to be Gaussian distributed in logarithmic scale. The following relative standard deviations have been applied (personal communication Michael Memmesheimer):

SO_2	$\ln(1.7) = 0.53$
NH_3	$\ln(1.7) = 0.53$
NO	$\ln(1.3) = 0.26$
others	$\ln(2.0) = 0.69$

Hence, for most emitted species a factor of 2.0 is assumed (this means that factors 2.0 and 0.5 define the error bar). Only SO_2, NH_3 and NO are defined to be better known.

The chemical correlation follows the ad hoc approach as described in Elbern and Schmidt (1999).

3.4 The minimisation procedure

The minimization procedure follows the forward model run, based on a prior or updated initial state and emission rates, followed by the adjoint (backward) integration. The input for the iterative minimisation procedure are the total costs $\mathcal{J}$, deviation $\delta\boldsymbol{x}$ from the background state $\boldsymbol{x}_b$, and the gradient $\partial\mathcal{J}/\partial(\delta\boldsymbol{x}_0(t_0), \delta\boldsymbol{u})^T$, due after each backward integration and resulting in an updated deviation from the background state and emission correction factor. Taking the first guess identical to the background state $\boldsymbol{x}_b$ and

$\boldsymbol{e}_b$ avoids the need to calculate any inverse of **B** and **K** nor square root thereof, and allows for the straightforward application of the pseudo-diffusion approach.

The following sequence performs the minimisation and associated preconditioning with transformations:

1. Calculate the transformed gradient (12) by executing vertical diffusion and horizontal diffusion, both with half the number of time steps as inferred from (18) and (19). Finally, the normalisation is applied.

2. Run minimisation routine with $\delta \boldsymbol{z}$ saved from the last minimisation output, (or $\delta \boldsymbol{z}$=0 for the first iteration).

3. Save $\delta \boldsymbol{z}$ for next iteration.

4. Apply transformation (15) to $\boldsymbol{v}$ to calculate $\delta \boldsymbol{x} = \mathbf{B}^{1/2}\boldsymbol{v}$.

5. Add first guess $\boldsymbol{x}_b$ to $\delta \boldsymbol{x}$ to obtain improved initial state values for the next iteration, and likewise for $\delta \boldsymbol{u}$.

The quasi–newton limited memory L-BFGS algorithm described in Nocedal (1980) and Liu and Nocedal (1989) is applied for the minimisation, after modification to a parallel version. The assimilation procedure is taken as successfully finished after the minimum is attained, while the a posteriori validation (see Sect. 5) is passed.

4 Observational data basis

4.1 Available data

Surface in situ observations assimilated for this study stem from the archives of the European Environmental agency (EEA) as compiled at the time of study (see AirBase – EEA's public European Air quality data base http://air-climate.eionet.europa.eu/databases/airbase/, and from some national and regional environmental agencies.

The EEA database covers geographically all countries from the European Union, the EEA member countries and some EEA candidate countries. EEA files contain information submitted by the participating countries throughout Europe. For some countries, for whom national and regional data sources with higher observation density were available, EEA data has been replaced by original sources. These include: the environmental protection agencies from the German states (LUA) and the federal state (UBA), Switzerland, Austria, The Netherlands, Denmark, and the United Kingdom. Measurement sites are operated routinely under the auspices of government authorities. A partly dense, partly very coarse coverage of Europe by observation sites is provided by this configuration (see Fig. 4). Including all types of stations, a typical day consists of about 130 000 observations, with about 13% observations of SO_2, 37% observations of NO_2, 20% observations of NO, and 30% observations of O_3. In most cases, NO_x data are available with a half–hourly frequency, while other data with hourly frequency.

The air quality database consists of air quality measurement data and their statistics for a representative selection of stations and for a number of pollutants, also meta-information on the involved monitoring networks, their stations, and measurments. Observations of SO_2, O_3, NO, NO_2, and CO are assimilated. Due to its small number, and therefore lack of statistical representativity, CO assimilation results are not presented in this study.

4.2 Observation error covariance matrix

In the cases presented here, observation errors are not provided with the data. Experiences from the BERLIOZ experiments are now used to determine the measurement and representativeness errors of the observation. The measurement error is chosen following a scheme exposed in Mohnen (1999), defining a relative error and a minimal absolute error for each species:

$$r_{\text{meas}} = \max\left(e^{\text{abs}}_{\text{min}}, e^{\text{rel}} \times y^{o}\right) \tag{25}$$

The representativeness of an observation is depending on the grid resolution and

the characteristics of the location. Most of the European organisations provide characteristics of their measurement locations. This information has been translated to this scheme.

The representativeness errors are encoded in the current study as a characteristic absolut error ϵ^{abs} (see Table 2) portion for each species, scaled by a factor depending on the grid resolution Δx and a characteristic representativeness length L_{repr}, given for the measurement station characteristics in Table 3:

$$r_{\text{repr}} = \sqrt{\frac{L_{\text{repr}}}{\Delta x}} \times \epsilon^{\text{abs}} \tag{26}$$

The observation error covariance matrix is here assumed to be diagonal, that is, the observation errors are not correlated. The diagonal elements of **R** are thus given by the sum of the measurement and representativeness errors,

$$R_{ii} = r_{i,\text{meas}} + r_{i,\text{repr}}, \quad i = 1, \ldots, n_{\text{obs}}. \tag{27}$$

5 Case study and results

5.1 Case study period and meteorological conditions

From 3 to 20 August 1997, a long lasting episode of elevated ozone levels over central Europe took place. The mesoscale meteorological simulations of this time span performed by MM5 has been restarted every 48 h, beginning with 1 August 00:00 GMT until 20 August 24:00 GMT. Meteorological initial and boundary values were taken from ECMWF analyses. The meteorological and chemistry–transport simulation of the first two days of the episode serve to attain a chemically balanced initial state by 3 August.

The ozone episode started with a shallow ridge with south-western to north-eastern tilt, extending from the Alps to southern Finland at 3 August. The centres of shallow adjacent depressions were situated south-west of the British isles in the west and over

the west coast of the Black Sea in the east. During the following days the ridge moved slightly westward, attaining a more north-south alignment. At 7 August, a flat Scandinavian high pressure system evolved from the ridge with a saddle point over central Europe. Regional weather conditions in that area were mostly sunny with local thunderstorms. After weakening of the high, a blocking ridge still prevailed with small variations of its longitudinal position until 17 August, when a shallow summerly depression developed over the river Rhine area. This low moved slightly eastbound while filling up in the flat but increasing blocking ridge until 20 August. Surface weather conditions were characterised by weak varying winds and occasional local thunderstorms over central Europe during these last days of the episode. After 21 August 1997, unsettled conditions resumed over central Europe and enforced the end of this episode with elevated ozone levels.

5.2 Assimilation runs set-up

To estimate the success gained by the 4D-var assimilation process, a suite of experiments has been conducted, which differ in various ways. The assimilation window is set to 24 h, from midnight to midnight, unless otherwise noted. The ensuing forecasted second day only serves for success control by improved predictions, stated by reduced model-minus-observation discrepancies. Hence, the study period consists of a sequence of assimilation days and forecast days. In each sequence, the background chemical state field is taken from the 2 day forecast, resting on the related assimilation result. Likewise, the new background emissions are taken from the preceding analysis result for the emission factors.

The case study comprises assimilation procedures in three different modes:

1. only initial value optimisation (IV),
2. only emission rate optimisation (ER), and
3. joined initial value/emission rate optimisation (IE).

All initial assimilation runs of the case study, that is those starting with 3 August after 2 days spin-up time, have been performed with background emission rates taken from the EMEP emission inventory. The only exception is an assimilation sequence in order to verify the analysed correction factors of the emission rates. In this case, a second suite of assimilation runs has been performed with significantly different emission rates, aiming to approximate the analysis result of the first suite. Further, the value of NO_x measurements is assessed by omitting these data, and comparing the analysis results with the NO_x-observation augmented case.

5.3 A Posteriori Validation

The validity of the assimilation results can only be shown satisfactorily by independent observations. Retaining observations from assimilation and preserving them for quality check is the usual way to demonstrate assimilation improvements. Given the task to ameliorate forecasts, an improved prediction skill also provides a means of validation, when chemical state forecasts with and without data assimilation are compared. Both modes are presented in this study.

Given properly defined background and observation error covariance matrices, and further given a statistically representative and unbiased innovation vector $\boldsymbol{d}=\boldsymbol{y}-\mathbf{H}\boldsymbol{x}$, the expected minimum $\mathcal{J}_{\min}$ of a quadratic objective function of the form (2) is equal to

$$\begin{aligned}\mathcal{E}(J_{\min}) &= \tfrac{1}{2}\mathcal{E}\left(\operatorname{trace}\left(\left(\mathbf{HBH}^T+\mathbf{R}\right)^{-1}\left(\boldsymbol{d}\boldsymbol{d}^T\right)\right)\right)\\ &= \tfrac{1}{2}\operatorname{trace}\left(\left(\mathbf{HBH}^T+\mathbf{R}\right)^{-1}\mathcal{E}\left(\boldsymbol{d}\boldsymbol{d}^T\right)\right)\\ &= \tfrac{1}{2}\operatorname{trace}\left(\mathbf{I}_M\right) = M/2, \end{aligned} \tag{28}$$

where $\mathbf{I}_M$ is the unit matrix in observation space. The generalisation for the spatio-temporal case is straightforward. Hence, it is easy to check whether this condition (often called the χ^2-condition) is satisfied (Talagrand, 1998, 2004; Ménard et al., 2000).

With the error estimates described in section 2.4 and averaged over all eight assimilation days, mean costs of 50×10^3 are inferred, with a halfed number of observations of 67×10^3 . This indicates a moderate overestimation of presumably observation errors, as both types of background costs remain too low to be effective in modifying the total costs. Rather, given the coarse resolution of the grid, the error of representativeness is suspected to contribute most to the deficit from the optimal cost value. In the case of nitrogen (di)oxide measurements, the high error of representativity attributed may be a specific reason. While modifying error estimates is simple in technical terms, only long and continuous operational applications, comprising different weather conditions, can give aid to better estimate the error statistics involved in the cost function.

As an example for a single assimilation run with joint initial value/emission rate optimisation, Fig. 5 exhibits iteration dependent decreases of cost results for 9 August 1997, with absolute costs broken down in terms of species and observation types. The minimisation of the partial costs of ozone, which contributes the highest portions in most cases, then virtually determines minimisation progress. Minimisation of SO_2 shows a similar reduction, however on a much smaller absolute basis. Further, NO_x species show nearly no minimisation effect. This is an indication that the 54 km grid cells are not suited to resolve emission patterns and surface fields of NO_x. The costs emerging from the iterative digression from background values of both initial values and emission rates, remain low. Nevertheless, the background term is indispensable for both reasons, to formally pose an overdetermined optimisation problem and practically to enforce an analysis result complying with model and observations.

5.4 Forecast improvements

The relative importance of emission rate optimisation compared to initial value optimisation must be expected to be highly dependent on conditions like emission strengths, boundary layer height and stability, and chemical life time. Therefore, in order to provide a representative example illustrating typical difficulties, local conditions should be selected which are influenced by both rural and urban conditions inside a model grid

box. With the dates 9 and 10 August 1997, this case is just in the middle of the then mature ozone episode. Figure 6 presents an example case, which is amidst the central European region, where the small to medium size cities of Wetzlar, Gießen, and Linden are surrounded by rural conditions. While the former stations are classified as sites "traffic", Linden is attributed the "rural" type. Nevertheless, visual inspection of the observed time series clearly indicates a strong diurnal cycle also for this station. It must be concluded, that the available measurements, at least for NO and NO_2, may be of limited representativity for the model grid cell.

In the case of SO_2, only a short measurement sequence of a few hours duration is available for each day. Observed values of about 3 to 7 ppbV are strongly overpredicted without any data assimilation, except during the last 12 h of the 2-day simulation cycle. Pure emission rate optimisation, building on inversion for the preceeding days 3, 5, and 7 August, is successful to predict the 10 August levels, after marginal overprediction of the assimilated measurements available during the morning hours of 9 August. Pure initial value optimisation is able to better simulate the assimilated observations. However, in contrast to emission rate optimisation, strong relaxation toward the simulation without data assimilation occurs in the course of the forecasted second day. For SO_2, it can be concluded that emission rates are more likely to be the right optimisation parameter. It can be corroborated by visual inspection of forecast improvements in this special grid cell, that combined emission rate-initial value optimisation does not provide for a better performance than in the case of pure emission rate optimisation.

It has been stated above, that the representativity of routinely operated NO and NO_2 observations in typical central European areas is poor. The vicinity of streets and the practice of environmental agencies to deploy observation sites mostly in populated areas renders assimilation of these data critical. In this study, this fact has been taken into account by increasing the error of representativity and hence the overall observation error as described in Sect. 3, thereby reducing the effect of assimilation of these observations. Consequently, deviations from the free simulation of all assimilation based simulated NO und NO_2 times series are hardly visible in Fig. 6. Further, simulated con-

centration levels of NO und NO_2 remain significantly lower than observed levels, but still inside the error margins. This is a direct consequence from the unequal distribution of observation sites, favouring densely habitated areas with enhanced emission levels.

Ozone data assimilation results for the selected sites show clearly a different performance. Observations exhibit distinct diurnal cycles with peak values between 60 and 70 ppbV during the afternoon hours and massive ozone depletion after midnight, which is a typical behaviour for strong anthropogenic emissions. The free simulation predicts too high concentration levels by 20 to 30 ppbV, exceeding error margins. Further, the observed strong nighttime concentration drop is only weakly featured by the model.

The data assimilation procedure with optimisation of only initial values shows a good agreement with observations of the first day, that is the assimilation window, including a much better, though not perfect simulation of the nighttime depletion. During the ensuing forecast of the second simulation day, the initial value optimisation based simulation slowly relaxes toward the free run, indicating the beneficial, yet short and ceasing impact of initial value modification, as was assessed in Elbern and Schmidt (2001).

The pure emission rate optimisation case shows nearly a coincidence with the free simulation during the first day, but significant improvements for the second, forecasted day. This demonstrates that (i) this specific model area has not yet benefitted from data assimilation of preceding days of the case study, and that (ii) emission optimisation of emitted precursor species requires a latency time prior to observing beneficial effects for photooxidant products like ozone. While obviously emission rates need to be optimised, the initial values of ozone remain poor.

Generally, for a satisfying inversion procedure, benefits from both procedures are expected. This can be observed from the case of combined emission rate-initial value optimisation. Here, the results approximate the initial value optimisation case during the first day and the emission rate optimisation case during the second day. In addition, the nightime ozone depletion is better forecasted than in either homogeneous case.

To present representative results, the relative benefits of data assimilation proce-

dures with and without emission rate optimisation will be discussed by examining biases and root mean square errors (RMS) of ensuing forecasts, resting on data assimilation in a variety of configurations.

For sulfur dioxide Fig. 7 presents mean bias and RMS time series for two day forecasts, averaged over eight consecutive 48 hour simulations from 3 to 18 August 1997. The first 24 h cover the assimilation interval. The extension to forecasts for the second day are displayed for quality control only.

Without any data assimilation the observation-minus-model bias (OmM) gives an average of 8 ppbV too high simulated values and a root mean square error (RMS) of about 10 ppbV. During the two day simulation the performance is poorer at hours centred around 07:00 UTC, which means 09:00 local summer time in most European countries, where observations were available. The possible reasons for this are still speculative, where poor vertical exchange due to a surface layer modelled too cold, poor vertical resolution, inexact diurnal emission profiles, and combinations thereof might be the reason. While these possible reasons point toward a violation of the "perfect model assumption" commonly made in 4D-var data asimilation, errors appear not to be strong enough to degrade later simulation severely, as the onset of vertical mixing during the following morning relaxes the problem. However, future investigation will focus on possible causes for the intermittent model performance drops at the surface layer.

Data assimilation with respect to optimisation of initial values only (IV case) shows a slight positive bias at initial time, which is a compensation for later relaxation toward the free simulation within the assimilation window. Later during the forecast period (that is, the second day), bias values approximates the reference run, providing a time scale for the memory for initial values in the system. A similar bevaviour is visible for the RMS error, which starts with an average of 5 ppbV and ends after 48 h with 15 ppbV only 1 ppbV less than that of the free simulation. From these error features it must be concluded that a severe bias in prevails. As mentioned above, this study adopts the hypothesis that one principal part of the bias stems from erroneous emission rates,

apart from initial values.

Pure emission rate optimisation results in a dramatical reduction of the bias down to values lower than 20%. Likewise, the RMS errors are reduced to levels between 30 and 50%. However, the initial RMS error of the initial value optimisation is smaller during the first five hours, clearly indicating an example of "over-optimisation" due to the optimisation of the less influential parameter, as this assimilation result is not sustainable over 48 h.

The data assimilation performance after joint optimisation of emission rates and initial values shows a further improvement for both bias and RMS error. It combines the improved features from initial value optimisation with those from emission rates, exhibiting the expected sustainability throughout the forecasted second day.

In a similar way, Fig. 8 presents biases and RMS errors for ozone in the control run, with values of about 10 ppbV and 22 ppbV, respectively. The early morning degradations are visible as for SO_2, indicating that the problem is not dominantly caused by chemical transformation. On the other hand, it is known that the coarse resolution of 54 km is not sufficient to resolve point and line sources of emitted ozone precursors correctly, leading to NO_x levels, which are too low to reduce ozone levels as observed at sites, mostly deployed close to NO_x emission sources.

Initial value optimisation reduces the bias mostly, inside the assimilation intervall a bias is even slightly reversed. A tendency for relaxation toward the free simulation can be observed for the second, forecasted day. As in the corresponding case of SO_2, the averaged RMS errors are reduced during the assimilation window, with later relaxation toward the free run average.

The pure emission optimisation based assimilation provides a substantial bias reduction of more than 50%, without attaining the success achieved for SO_2 however. In contrast, the RMS average values are only marginally better than exhibited for the free run. It can be concluded that emission rate optimisation gives a smooth control over the concentration levels, while initial value optimisation accounts for smaller scale variations, which are more suitable to fit to observed temporal variations.

The combined emission rate and initial value optimisation gives the best results, however only with modest improvements in comparison with the initial value optimisation. The assimilation interval bias is smallest and there is less relaxation of the bias toward the free run during the final hours of the forecasted second day. Also RMS errors are only slightly smaller than in the case of initial value optimisation.

While there is a clearly visible benefit for optimising both initial value and emission rates of precursors for ozone jointly, there is only a moderate improvement with respect to initial value only optimisation. The reason for this is presumably the coarse resolution and the effect, that gross biases of emission rates are reduced during the first two-day cycles of the case study, leaving the remaining discrepancies to initial value optimisation.

Validation of results can also be undertaken by observations withheld from the assimilation algorithm, in addition to assessing increments in prediction skills after data assimilation. In the sequel, a combination of both methods is presented, where observations of quasi-randomly selected measurement sites were not assimilated, but taken for validation of assimilation and forecast results. Figure 9 presents the average results for SO_2 and O_3, again by bias and RMS. For clarity only the joint emission rate-initial value optimisation option is selected. The free simulation results of the stations withheld are included for comparison only.

In addition to the initial value/emission rate optimisation based simulation, the first guess based simulation is included, to assess possible accumulated benefits from earlier data assimilation runs, made in two days steps. It can be seen that there is a drastic improvement for SO_2, with nearly no distinction from the actual assimilation based simulation. In contrast, for ozone a significant improvement for the performance at the locations of withheld stations can be claimed, when compared with the free run. Improvements with respect to the first guess based line indicate information gain achieved from the other 75% observation sites.

In comparison with ozone and its precursors, it can be concluded that the success of SO_2 forecasts after assimilation is due to the better representativity of grid cells,

less temporal and spatial variability, and considerably reduced chemical dynamics on time scales of days. Consequently, a denser observational network is required as well, to significantly improve performance in the realm of tropospheric photochemistry data assimilation. Further, basic assumptions of 4D-var, like the tangent-linear approximation being sufficient and prevalence of Gaussian error characteristics must be rendered moot for this case. Nevertheless, both forecast improvements and analysis improvements at stations withheld from assimilation prove a benefitial impact from the assimilation procedure.

5.5 Emission optimisation results

Besides improved forecast skills, improved estimates of emission rates can be expected from the 4D-var emission rate inversion procedure. Correction factors should be, with variable degree of confidence, provided for all 19 emitted species. Formally, the 4D-var procedures performed over the 16 days result in independent analyses of emission rates. Despite its differentiation in terms of working days, saturdays and sundays, there is surely an unknown day-to-day variability in emissions, which the emission inventory may not fully capture. Nevertheless, under conditions described here, and taking note of the results described above, the emission inventory must be assumed to be biased, with respect to real emission rates including its margins of daily variations. The possible presence of biases has been ignored in the assimilation procedure. At least in cases of not too strong differences it can be expected that the inversion procedures converge toward less biased emission estimates, which may be manifest in stabilising correction factors of emission rates. After inspection of the results, for the entire period the correction factors of the emission rates exhibit a stable tendency.

Figure 10 displays the analysed correction factors for the emission rates of sulfur dioxide, NO_2, terminal alkenes, and isoprene for the lowest model layer, based on joint emission rate-initial values optimisation with a 24 h assimilation interval. 17 August 1997 is the last day of the case study with data assimilation. As most of the integration domain is void of observations, interpretation is restriced to observed areas as

displayed in Fig. 4. In the case of SO_2 a general reduction by about 20 to 25% can be observed for England and the central European area, except in occasional locations. For the limited territory of the former German Democratic Republic, it can be concluded that the transition from older coal fueling power houses to fewer, but cleaner plants progressed more rapidly than estimated by the emission inventory. The correction factors for the Iberian peninsula are, not uniform, yet mostly amplifying the emissions.

A similar picture as for SO_2, though with less spread, can be stated for NO_2. In most observed areas, emission inventory rates must be reduced by a percentage of 15 to 20%. Nevertheless, various urban centres can be identified where a small increment up to 15% is inferred.

As examples for VOCs Fig. 4 presents correction factors for terminal alkene and isoprene. While the latter is purely emitted by deciduous forests, the former includes anthropogenically emitted alkenes and terpenes from coniferous forests. In either case, and in contrast to SO_2 and NO_2, there are major contiguous regions in eastern central Europe, where the emission parameterisation proved to provide slightly too low values of about 10 to 15%. Probably, this effect is due to an underestimation of biogenic VOC emissions. Generally, for VOC emission rate optimisation correction factors excert a less significant deviation from unity.

As SO_2 and NO_2 emission rates from the EMEP emission inventory proved to be too high, the result comes under scrutiny by starting with significantly too low emission rates, expecting the result to converge toward the same analysed absolute emission rates. Therefore, validation of emission correction factors inferrerd from EMEP based initial background emission rates are contrasted with an analog inversion procedure, based on halfed initial background emission rates at the beginning of the case study.

Figure 11 exhibits the results for 17 August 1997, the same day as in the exposition before, with initial emission rates reduced to 50%. As expected, most areas with former drastic reduction are now characterised by amplified emission rates. For SO_2, mostly moderately amplifying factors are visible, and still reducing factors in eastern Germany, pointing toward a drastic overestimation of emission in that area by EMEP.

This corroborates the former result, again indicating a faster reduction of lignite coal combustion and more efficient SO_2 filter techniques, following the change of economic system after 1990 in that area.

In the case of NO_2, stronger amplification factors are analysed over central Europe, reflecting a less strong misspecification. Again, the approximation of emission rates from below converges to an absolute result similar to that from above in central Europe and England.

For terminal alkenes the correction factors remain small and without clear correction signal. Given an analysis virtually resting on ozone observations, and NO_x observations strongly devaluated by the error of representativity, a plausible and likely explanation for the NO_2 and VOC correction scenario can be explained in the frame of the Empirical Kinetics Modelling Approach (EKMA). See for example Kinosian (1982) for further explanations. It is known that the coarse spatial resolution of the model and the strict confinement of NO_x sources to point and line sources renders the system to be biased toward the preference of "NO_x constrained" states, where ozone formation is very sensitive to variations in NO_x emission rates and concentration levels, and rather insensitive to VOC changes. This implies, that under "NO_x constrained" conditions, and without VOC observations available, the optimisation system is too ill-posed to provide credible VOC correction factors.

In fact, the minimal impact gained from assimilation of NO_2 observations can be concluded from Fig. 12, where a very similar pattern and absolute values for reduction factors are exhibited as in Fig. 10. A closer inspection of earlier assimilated days corroborates this finding.

6 Conclusions

In the chosen frame of continental scale air quality modelling, the 4D–var method for inverse modelling of emission estimates has been shown to perform excellent for SO_2. This species may however serve as representative for slowly reactive species, with

emission patterns other than low level line sources like streets. In fact, for the European areas featured by the available observation sites, sulfur emission sources are characterised by the predominance of single point sources of power plants, rather than by a dense distribution of small sources from, say, housekeeping. In this context, the presented 4D-var approach has been able to demonstrate the outstanding importance of emission rate optimisation rather than initial value optimisation. If only initial values were optimised, estimated initial states relax toward a mean 10 ppb SO2 bias already after two days simulation, which is emission inventory induced. In contrast, emission rate optimisation reduces biases nearly perfectly and reduces RMS errors by about 60%, which may delineate the observational and representativity accuracy of the set-up of model resolution and observational network. It could be analysed that the emission inventory grossly overestimates sulfur emissions, especially in areas of economic transition in Germany.

The optimisation of emission rates in the case of photolytically active species proved to be considerably more challenging. With ozone being a central constituent of interest, the following conditions appear to exert a high influence on the predictive skill:

- with 54 km horizontal resolution, the feasible model grid resolution for continental scale regional model is too coarse in urban and densely populated rural areas with typical point and line source emission patterns,
- the observational network density of NO_x as emitted precursor species is biased toward populated regions, measuring higher NO_x levels than simulated on the coarse domain and requiring elevated representativity errors of the in situ measurements,
- in urban areas, simulated conditions suffer from a model proclivity for NO_x controlled conditions for ozone production, rendering emission estimates of VOC uncertain.

The following can be concluded: Despite the fact that, in most cases, observed NO_x levels are significantly higher than modelled NO_x levels, emission estimates indicate

markedly lower emission rates, to achieve better ozone forecasts. The precision of analysed reductions of NO_x emission may be hampered by the above mentioned coarse model grid. However, in view of two additional assimilation validation exercises, taking either emission rates halfed or without NO_x observations, the general tendency must be considered as correct.

This statement is corroborated by performance assessments of forecast skill, where the bias could be decreased and the RMS error remains lower than for a forecast without data assimilation longer than 48 h.

On the basis of this relative performance increment it can be expected that further difficulties for forecast improvements is most likely caused mainly by the coarse model resolution. Accepting the computational costs by selecting finer grids is the direct measure to avoid the double problem of nearly invaluable NO_x observations and the model proclivity to simulate NO_x constrained photochemical scenarios. Future work will therefore focus on substantial grid refinement, with 4d-var only feasible after introducing nesting techniques. Further to this, preconditioning by empirical factors, now optimised for the total model domain, must be refined to local conditions. It can be expected, that these measures offer further avenues for forecast skill improvements.

Acknowledgements. Personal thanks are due to M. Memmesheimer, RIU, for thorough discussions on emission rate modelling. The following institutions are gratefully acknowledged: EMEP (co-operative programme for monitoring and evaluation of the long range transmission of air pollutants in Europe), Norway, for the emission inventory, the European Environmental agency (EEA) for providing AirBase data. Meteorological analyses for the initial values of the MM5 were obtained from ECMWF. We also thank J. Nocedal, North-Western University, for making the L-BFGS code available. The adjoint code of the dynamic operators was developed by use of the AMC software, version 3.88, obtained from Ralf Giering, MPI of Meteorology, Hamburg, and by the Odyssee software, provided by C. Faure and Y. Papegay, INRIA, S. Antipolis, France. Access given to these software packages is highly appreciated. We gratefully acknowledge the computational support by the Central Institute for Applied Mathematics (ZAM) of the Research Centre Jülich and the Regional Computing Centre of the University of Cologne (RRZK). Financial support was provided by the European Union's Climate and Environmental

program project “Regional dIFferences in Tropospheric Ozone” (RIFTOZ) under grant ENV4–CT95–0025, and by the Atmosphärenforschungsprogramme AFO 2000 project “Tracer flux analyses of tropospheric field campaigns by chemical 4D-var data assimilation” (SATEC4D) of the German Ministry for Education, Science, Research, and Technology (BMBF) under grant 07AFT03.

References

Anthes, R. A., Hsie, E.-Y., and Kuo, Y.-H.: Description of the Penn State/NCAR Mesoscale Model Version 4 (MM4), Tech. Note NCAR/TN/282+STR, National Center of Atmospheric Research, Boulder, Colorado, 1987. 1737

Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: Theory, Q. J. R. Meteorol. Soc., 131, 2191–2208, 2005a. 1729

Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy. II: Applications, Q. J. R. Meteorol. Soc., 131, 2209–2223, 2005b. 1729

Bott, A.: A positive definit advection scheme obtained by nonlinear renormalization of the advective fluxes, Month. Weath. Rev., 117, 1006–1015, 1989. 1735

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual CO2 sources and sinks, 1. Method and control inversion, J. Geophys. Res., 104, 26 161–26 178, 1999a. 1729

Bousquet, P., Ciais, P., Peylin, P., Ramonet, M., and Monfray, P.: Inverse modeling of annual CO2 sources and sinks, 2. Sensitivity study, J. Geophys. Res., 104, 26 179–26 193, 1999b. 1729

Briggs, G. A.: Lectures on air pollutants and environmental impact analysis, chap. Plume rise predictions, pp. 59–111, Am. Meteorol. Soc., Boston, 1975. 1736

Chai, T. F., Carmichael, G. R., Sandu, A., Tang, Y. H., and Daescu, D. N.: Chemical data assimilation of Transport and Chemical Evolution over the Pacific (TRACE-P) aircraft measurements, J. Geophys. Res., 111, D02301, doi:10.1029?2005JD005883, 2006. 1728

Chang, J. S., Brost, R. A., Isaksen, I. S. A., Madronich, S., Middleton, P., Stockwell, W. R., and Walcek, C. J.: A three-dimensional Eulerian acid deposition model: physical concepts and formulation, J. Geophys. Res., 92, 618–700, 1987. 1735

Chao, W. C. and Chang, L. P.: Development of a four-dimensional variational analysis system

using the adjoint method at GLA Part I: Dynamics, Month. Weath. Rev., 120, 1661–1673, 1992. 1737

Courtier, P.: Dual formulation of four-dimensional variational assimilation, Q. J. R. Meteorol. Soc., 123, 2449–2461, 1997. 1739

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. 1732

Daley, R.: Atmospheric Data Analysis, Cambridge Univ. Press, 1991. 1731

Elbern, H. and Schmidt, H.: A four–dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, J. Geophys. Res., 104, 18 583–18 598, 1999. 1728, 1745

Elbern, H. and Schmidt, H.: Ozone Episode Analysis by Four-Dimensional Variational Chemistry Data Assimilation, J. Geophys. Res., D4, 3569–3590, 2001. 1728, 1733, 1735, 1739, 1753

Elbern, H. and Schmidt, H.: Chemical 4D variational data assimilation and its numerical implications for case study analyses, in: IMA volumes in Mathematics and its Applications, Atmospheric Modeling, edited by: Chock, D. P. and Carmichael, G. R., 130, 165–184, 2002. 1730

Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modeling, J. Geophys. Res., 102, 15 967–15 985, 1997. 1728

Elbern, H., Schmidt, H., Talagrand, O., and Ebel, A.: 4D–variational data assimilation with an adjoint air quality model for emission analysis, Environ. Model. and Software, 15, 539–548, 2000. 1730, 1732

Elbern, H., Baier, F., Bittner, M., Bochorishvili, R., Bovensmann, H., Hoffmann, L., Joppich, W., Meyer, J., Riese, M., Schwinger, J., Stiller, G., and von Clarmann, T.: SACADA, a new chemical data assimilation system for earth observation, in: Results of the German atmospheric research programme AFO2000, final publication, 210–216, GSF, 2005. 1728

Engelen, R., Andersson, E., Chevallier, F., Hollingsworth, A., Matricardi, M., McNally, A. P., Thépaut, J. N., and Watts, P. D.: Estimating atmospheric CO2 from advanced infrared satellite radiances within an operational 4D-Var data assimilation system: Methodology and first results, J. Geophys. Res., 109, D19309, doi:10.1029/2004JD004777, 2004. 1729

Enting, I. G. and Newsam, G. N.: Inverse problems in atmospheric constituent studies: II. Sources in the free atmosphere, Inverse Problems, 6, 349–362, 1990. 1729

Enting, I. G., Trudinger, C. M., and Francey, R. J.: A synthesis inversion of the concentration and d13C of atmospheric CO2, Tellus, 47B, 35–52, 1995. 1729

Errera, Q. and Fonteyn, D.: Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, J. Geophys. Res., 2001. 1728

Eskes, H. J., Piters, A. J. M., Levelt, P. F., Allart, M. A. F., and Kelder, H. M.: Variational assimilation of total-column ozone satellite data in a 2D lat-lon tracer-transport model, J. Atmosph. Sci., 56, 3560–3572, 1999. 1727

Fan, S. M., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., and Tans, P.: A large terrestrial carbon sink in North America implied by atmospheric data and oceanic carbon dioxide data and models, Science, 282, 442–446, 1998. 1729

Faure, C. and Papegay, Y.: Odyssée Version 1.7 Users Guide, Technical Report 0224, Institut National de Recherche en Informatique et en Automatique, manual adjoint compiler, 1998. 1737

Fedorov, V.: Kriging and other estimators of spatial field characteristics (with special reference to environmental studies), Atmos. Environ., 23, 174–184, 1998. 1727

Fisher, M. and Lary, D.: Lagrangian four-dimensional variational data assimilation of chemical species, Q. J. R. Meteorol. Soc., 121, 1681–1704, 1995. 1727, 1743

Giering, R. and Kaminski, T.: Recipes for Adjoint Code Construction, ACM Trans. Mathematical Software, 24, 437–474, adjoint code, 1998. 1737

Gloor, M., Fan, S. M., Pacala, S., Sarmiento, J., and Ramonet, M.: A model-based evaluation of inversions of atmospheric transport, using annual mean mixing ratios, as a tool to monitor fluxes of nonreactive trace substances like CO2 on a continental scale, J. Geophys. Res., 104, 14 245–14 260, 1999. 1729

Grell, G. A., Dudhia, J., and Stauffer, D. R.: A description of the fifth generation Penn State/NCAR mesoscale model (MM5); NCAR TN-398+IA, Tech. rep., National Center for Atmospheric Research, Boulder, Colorado, 1993. 1737

Gurney, K. R.: Towards robust regional estimates of CO2 sources and sinks using atmospheric transport models, Nature, 415, 626–630, 2002. 1729

Hanea, R. G., Velders, G. J. M., and Heemink, A.: Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport, J. Geophys. Res., 109, D10302, doi:10.1029/2003JD004283, 2004. 1730

Hanna, S. R., Chang, J. C., and Fernau, M. E.: Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables,

Atmos. Environ., 32, 3619–3628, 1998. 1728
Hanna, S. R., Lu, Z. G., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., and Hansen, D. A.: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July1995 OTAG domain, Atmos. Environ., 35, 891–903, 2001. 1728
Hass, H., Jakobs, H. J., and Memmesheimer, M.: Analysis of a regional model (EURAD) near surface gas concentration predictions using observations from networks, Met. Atmos. Phys., 57, 173–200, 1995. 1735
Hesstvedt, E., Hov, Ø., and Isaksen, I. S. A.: Quasi-steady state approximation in air pollution modeling: Comparison of two numerical schemes for oxidant prediction, Int. J. Chem. Kinet., 10, 971–994, 1978. 1736
Hölzemann, J., Elbern, H., and Ebel, A.: PSAS and 4D-var data assimilation for chemical state analysis by urban and rural observation sites, Phys. Chem. Earth, 26, 807–812, 2001. 1739
Houweling, S., Kaminski, T., Dentener, F., Lelieveld, J., and Heimann, M.: Inverse modeling of methane sources and sinks using the adjoint of a global transport model, J. Geophys. Res., 104, 26 137–26 160, 1999. 1729
Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C.: Unified notation for data assimilation: operational, sequential and variational, J. Meteorol. Soc. Japan, 75, 181–189, 1997. 1732
Issartel, J. P.: Rebuilding sources of linear tracers after atmospheric concentration measurements, Atmos. Chem. Phys., 3, 2111–2125, 2003. 1729
Jakobs, H. J., Feldmann, H., Hass, H., and Memmesheimer, M.: The Use of Nested Models for Air Pollution Studies: An Application of the EURAD Model to a SANA Episode, J. Appl. Meteorol., 34, 1301–1319, 1995. 1738
Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport: 1. Adjoint model and Jacobian matrix, J. Geophys. Res., 104, 18 535–18 553, 1999a. 1729
Kaminski, T., Heimann, M., and Giering, R.: A coarse grid three-dimensional global inverse model of the atmospheric transport: 2. Inversion of the transport of CO2 in the 1980s, J. Geophys. Res., 104, 18 555–18 581, 1999b. 1729
Kaminski, T., Knorr, W., Rayner, P. J., and Heimann, M.: Assimilating atmospheric data into a terrestrial biosphere model: A case study of the seasonal cycle, Global Biogeochem. Cycles, 16, 1066, doi:10.1029/2001GB001463, 2002. 1729
Kinosian, J. R.: Ozone-precursor relationships from EKMA diagrams, Environ. Sci. Technol.,

16, 880–883, 1982. 1759
Liu, D. C. and Nocedal, J.: On the limited memory BFGS method for large scale optimization, Math. Programming, 45, 503–528, 1989. 1746
Lorenc, A. C.: Analysis methods for numerical weather prediction, Q. J. R. Meteorol. Soc., 112, 1177–1194, 1986. 1731
Lorenc, A. C.: Optimal nonlinear objective analysis, Q. J. R. Meteorol. Soc., 114, 205–240, 1988. 1731
Madronich, S.: Photodissociation in the atmosphere. 1. actinic flux and the effects of ground reflections and clouds, J. Geophys. Res., 92, 9740–9752, 1987. 1736
McRae, G. J., Goodin, W. R., and Seinfeld, J. H.: Numerical solution of the atmospheric diffusion equation for chemically reacting flows, J. Comp. Phys., 45, 1–42, 1982. 1735, 1736
Memmesheimer, M., Hass, H., Tippke, J., and Ebel, A.: Regional Photochemical Measurement and Modeling Studies, vol. 2, chap. Modeling of episodic emission data for Europe with the EURAD Emission Model (EEM), pp. 495–499, Air & Waste Management Association, Pittsburgh, USA, 1995. 1737
Ménard, R., Cohn, S. E., Chang, L.-P., and Lyster, P. M.: Assimilation of Stratospheric Chemical Tracer Observations Using a Kalman Filter, Part I: Formulation, Month. Weath. Rev., 128, 2654–2671, 2000. 1750
Mohnen, V.: Data Quality Assessment – an overview for the TRACT 16./17. September 1992 field intensive, Report to the coordinator of TFS-LT1, IFU Garmisch-Partenkirchen, error estimation, data quality, 1999. 1747
Muller, J. F. and Stavrakou, T.: Inversion of CO and NOx emissions using the adjoint of the IMAGES model, Atmos. Chem. Phys., 5, 1157–1186, 2005. 1730
Navon, I. M.: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dyn. Atmos. Oceans, 27, 55–79, 1997. 1731
Newsam, G. N. and Enting, I. G.: Inverse problems in atmospheric constituent studies: I. Determination of surface sources under a diffusive transport approximation, Inverse Problems, 4, 1037–1054, 1988. 1729
Nocedal, J.: Updating Quasi-Newton Matrices With Limited Storage, Math. Comput., 35, 773–782, 1980. 1746
Quélo, D., Mallet, V., and Sportisse, B.: Inverse Modeling of NOx Emissions at Regional Scale over Northern France. Preliminary Investigation of the Second-Order Sensitivity, J. Geophys. Res., 110, D24310, doi:10.1029/2005JD006151, 2005. 1730

Robertson, L. and Langner, J.: Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment, Atmos. Environ., 32, 4219–4225, 1992. 1729
Rostaing, N., Dalmas, S., and Galligo, A.: Automatic differentiation in Odyssee, Tellus, 1993. 1737
Schmidt, H. and Martin, D.: Adjoint sensitivity of episodic ozone in the Paris area to emissions on the continental scale, J. Geophys. Res., 108, 8561, doi:10.1029/2001JD00158, 2003. 1728
Stajner, I., Riishøjgaard, L. P., and Rood, R. B.: The GEOS ozone data assimilation system: Specification of error statistics, Q. J. R. Meteorol. Soc., 127, 1069–1094, 2001. 1727
Stockwell, W. R., Middleton, P., and Chang, J. S.: The second generation regional acid deposition model chemical mechanism for regional air quality modeling, J. Geophys. Res., 95, 16 343–16 367, 1990. 1728
Struthers, H., Brugge, R., Lahoz, W. A., O'Neill, A., and Swinbank, R.: Assimilation of Ozone Profiles and Total Column Measurements into a General Circulation Model, J. Geophys. Res., 107, 4438, doi:10.1029/2001JD000957, 2002. 1727
Talagrand, O.: A posteriori evaluation and verification of analysis and assimilation algorithms, in: Proceedings of the Workshop on Diagnosis of Data Assimilation Systems, European Centre for Medium-range Weather Forecasts Reading, England, 2–4 November, 1998. 1741, 1750
Talagrand, O.: Objective Validation and Evaluation of Data Assimilation, in: Proceedings of the Seminar on Recent developments in data assimilation for atmosphere and ocean, ECMWF, European Centre for Medium-range Weather Forecasts, 2004. 1750
Talagrand, O. and Courtier, P.: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, Q. J. R. Meteorol. Soc., 113, 1311–1328, 1987. 1731
van Loon, M., Builtjes, P. J. H., and Segers, A. J.: Data assimilation of ozone in the atmospheric transport chemistry model LOTOS, Environ. Model. and Software, 15, 603–609, 2000. 1730
Verlaan, M. and Heemink, A. W.: Reduced rank square root filters for large scale data assimilation problems, in: Second International Symposium on Assimilation of Observations in Meteorology and Oceanography, 1995. 1730
Weaver, A. and Courtier, P.: Correlation Modelling on the Sphere Using a Generalized Diffusion Equation, Q. J. R. Meteorol. Soc., 127, 1815–1846, data assimilation covariance matrix, 2001. 1740, 1743
Wesely, M. L.: Parameterization of surface resistances to gaseous dry deposition in regional-

scale numerical models, Atmos. Environ., 23, 1293–1304, 1989. 1737
Yanenko, N. N.: The method of fractional steps: solution of problems of mathematical physics in several variables, Springer, 1971. 1735

Table 1. Selected minimal background errors $\epsilon_{abs}(l)$ for species in the RADM mechanism.

SO_2	H_2SO_4	O_3	NH_3
20 ppb	5 ppb	15 ppb	2 ppb
CO	HCHO, TOL	NO_x	others
800 ppb	10 ppb	5 ppb	1 ppb

Table 2. Charateristic error portion used to calculate the representativeness error for a ground based observation.

Species	ϵ^{abs} [ppbV]
SO_2	0.4
NO_2	1.4
NO	3.0
O_3	1.2
CO	15.0

Table 3. Radii of influence associated with each type of ground based in-situ observation.

Station Type	L_{repr} [km]
Remote	20.0
Rural	10.0
Suburban	4.0
Urban	2.0
Traffic	1.0
Unknown	3.0

ACPD
7, 1725–1783, 2007

EGU

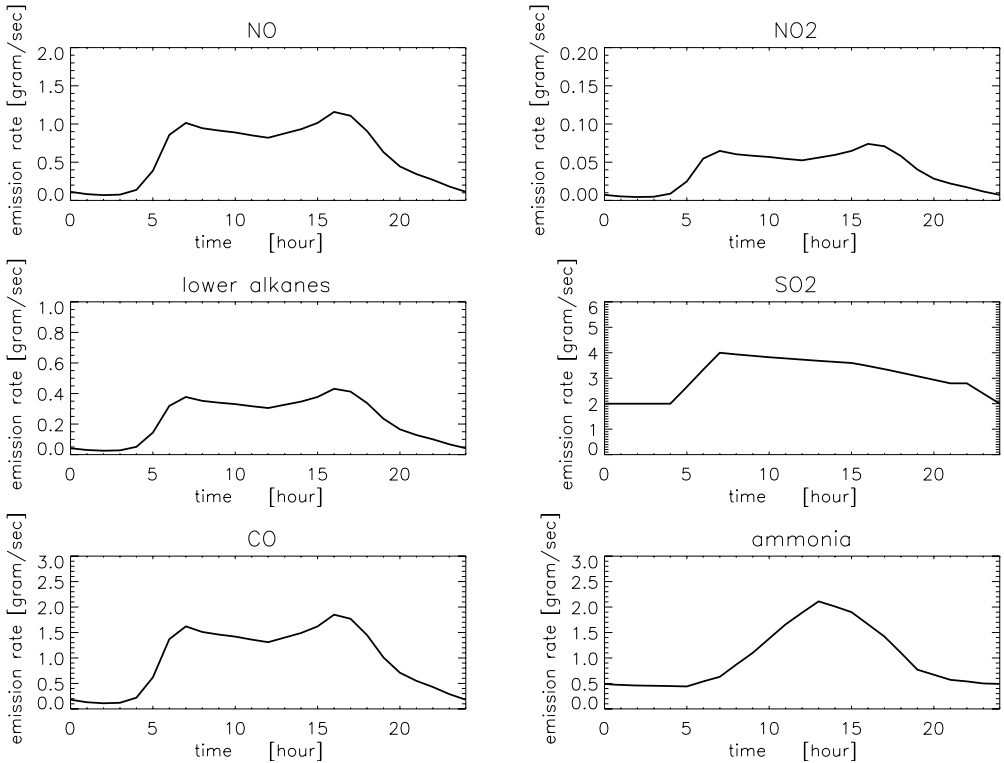

Fig. 1. Example diurnal profiles applied in the emission module for six of the 19 emitted species. Given are profiles for working day conditions of NO, NO_2, lower alkanes, SO_2, CO, and ammonia, as indicated.

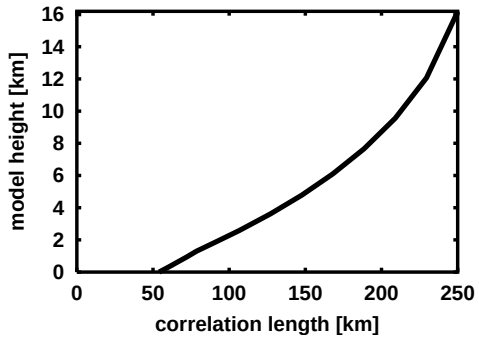

Fig. 2. Horizontal influence radius as a function of model height.

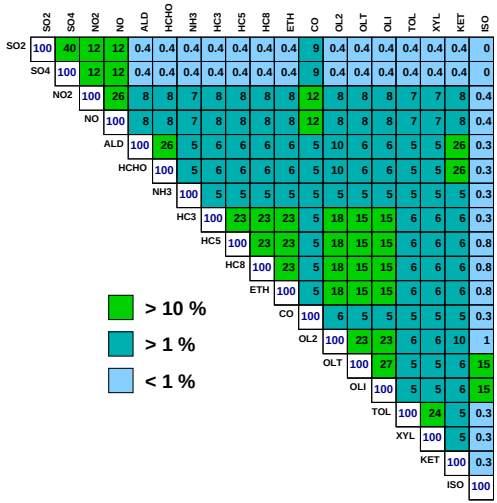

Fig. 3. Implemented background emission rate correlation matrix. All correlation values are given in percent.

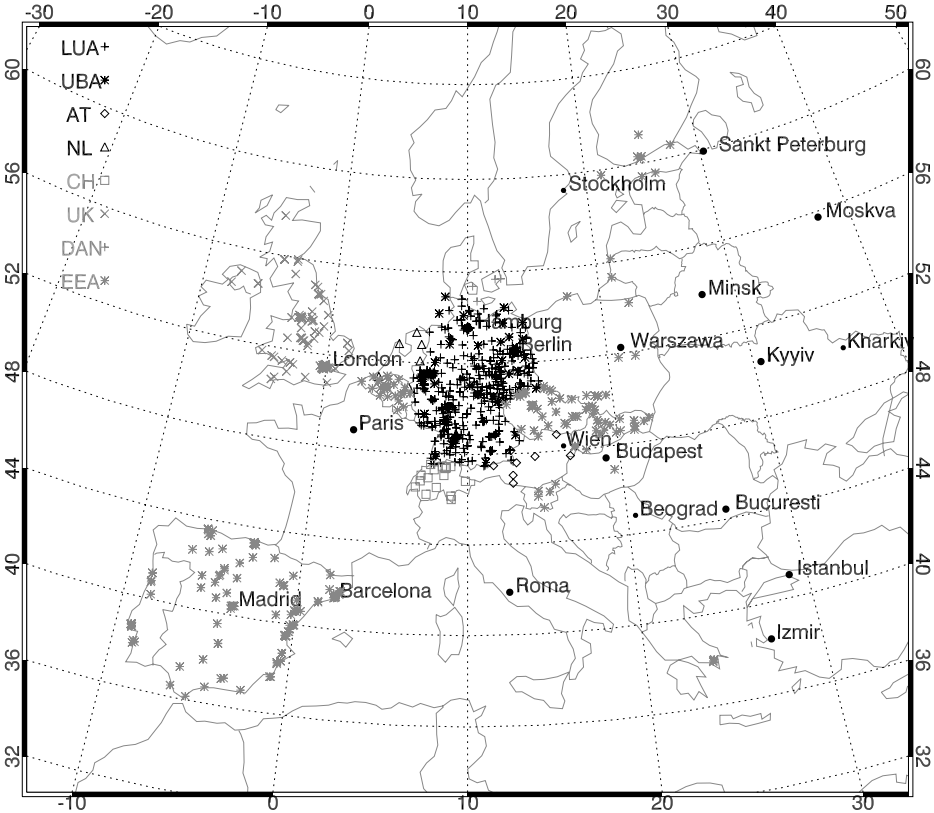

Fig. 4. The EURAD integration domain resolved with 54 km resolution and locations of available surface measurement stations for 9 August 1997. Contributions from the European Environmental Agency (EEA, grey ⋆), environmental protection agencies from the German states (LUA, black +) and the federal state (UBA, black ⋆), Switzerland (CH, grey □), Austria (AT, black ◊), The Netherlands (NL, black △), Denmark (DAN, grey +), and the United Kingdom (UK, grey ×).

ACPD
7, 1725–1783, 2007

EGU

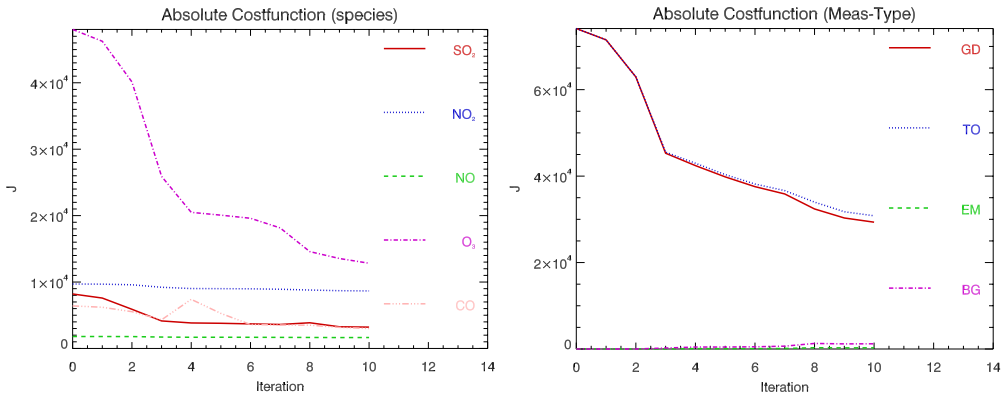

Fig. 5. Iterative decrease of partial and total costs of the joint emission rate/initial value optimisation for the 24 h assimilation interval of 9 August 1997. Top panel with partial costs of observed species. Bottom panel with cost types, where GD: surface in situ observations, BG: background costs of initial values, EM: background costs of emission rates, TO: total costs.

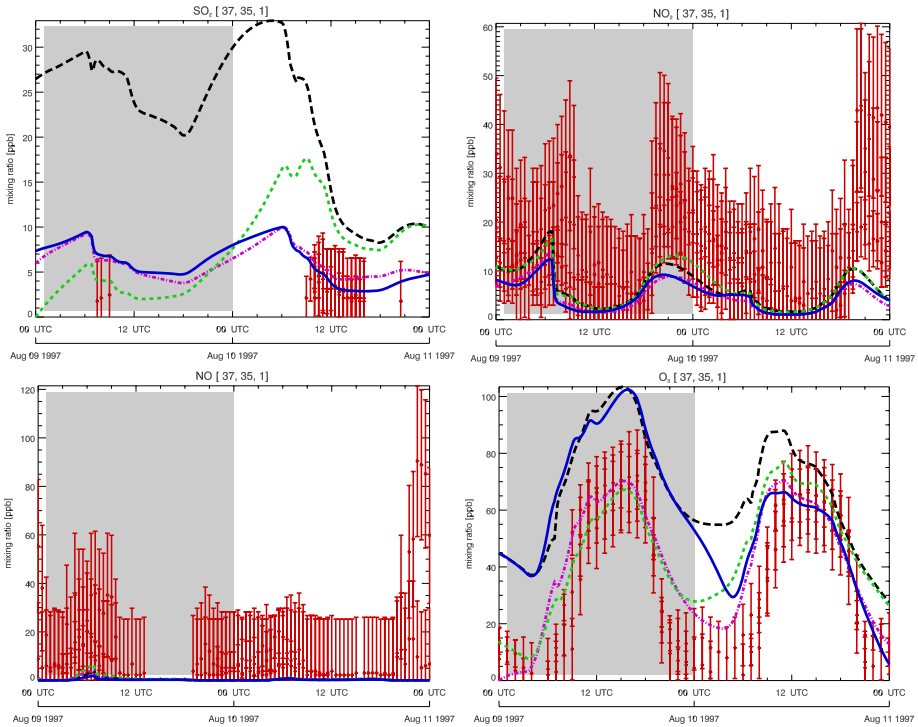

Fig. 6. Modelled and observed 48 h time series starting 9 August 00:00 UTC for a central European model grid box covering the 3 urban influenced measurement stations Gießen, Wetzlar, and Linden, with values given by red dots with error bars. The assimilation window covers the first 24 h within the grey shaded area. Later observations of 10 August serve for quality control only. Top left panel SO_2, top right: NO_2, bottom left NO, and bottom right O_3. Green dotted line: initial value optimisation, blue solid line: emission rate optimisation, pink dash-dotted line: combined initial value-emission rate optimisation. For comparison black dahed line: run without any data assimilation on this and prior days, started with spin-up background values of the spin-up period 1–2 August.

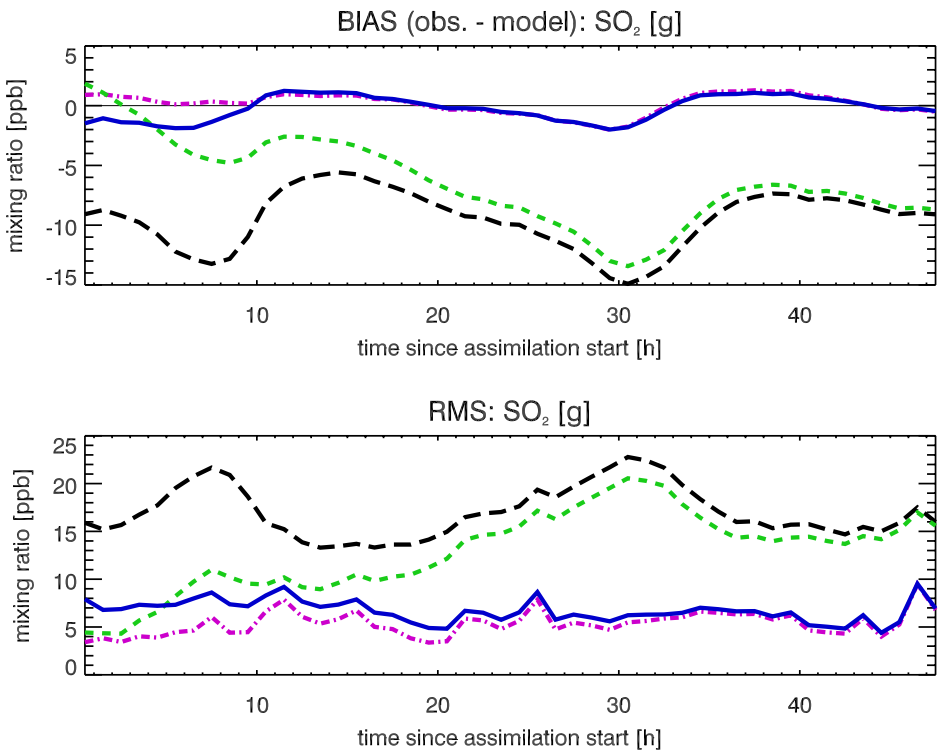

Fig. 7. Mean observation-minus-model differences (top) and RMS differences (bottom) of all 8 two-day simulations for sulfur dioxide. Observations assimilated the first 24 h, with hours 25–48 displayed for quality control only. Green stippled line: only initial value optimisation, blue full line: only emission rate optimisation, pink dash-dotted line: combined initial value/emssion rate optimisation. Dashed line: control run without any data assimilation in the case study, for reference only.

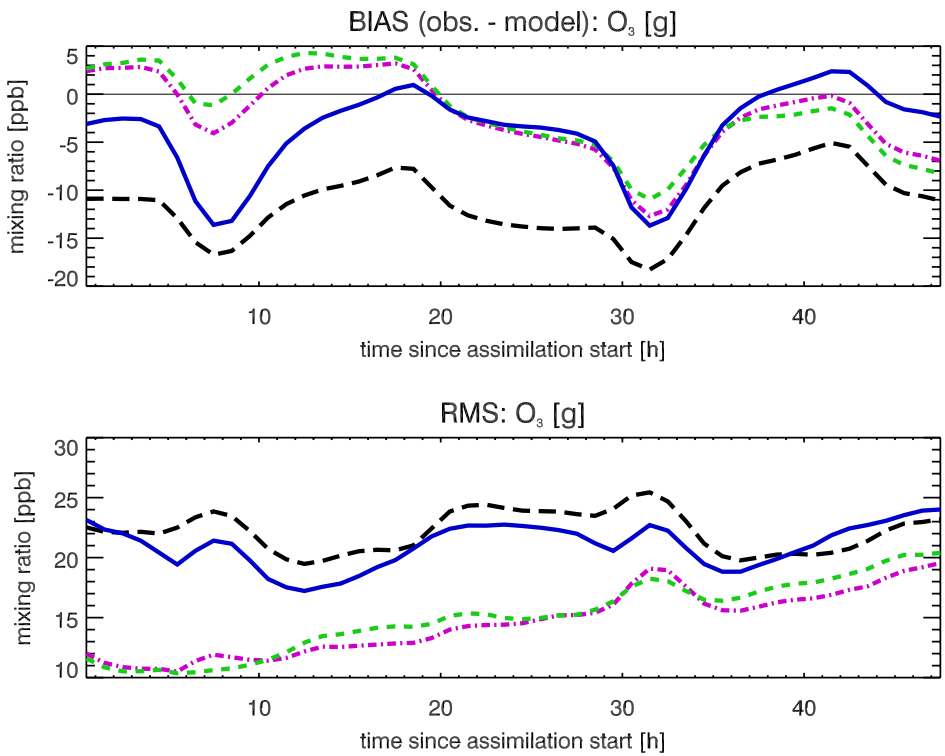

Fig. 8. Plotting conventions as for Fig. 7, except for ozone.

ACPD
7, 1725–1783, 2007

EGU

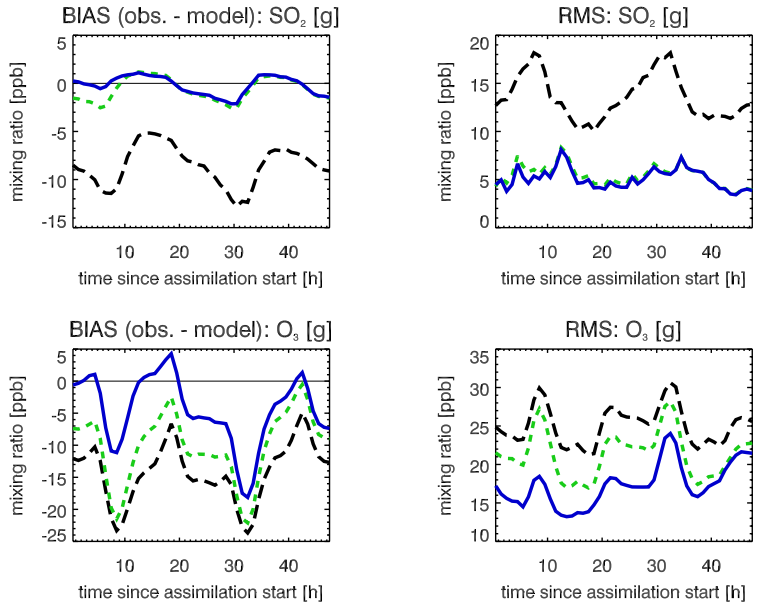

Fig. 9. Mean observation-minus-model differences (left panels) and RMS differences (right panels) of all 8 two-day simulations for sulfur dioxide (top panels) and ozone (bottom panels) for 25% randomly selected observation sites. Green stippled line: first guess performance at locations at randomly selected sites, the data of which is not assimilated. Blue full line: performance at selected sites, withheld from assimilation. Dashed line: control run without any data assimilation in the case study, for reference only. Observations assimilated the first 24 h in the joint initial value/emission rate optimisation mode, with hours 25–48 displayed for quality control only.

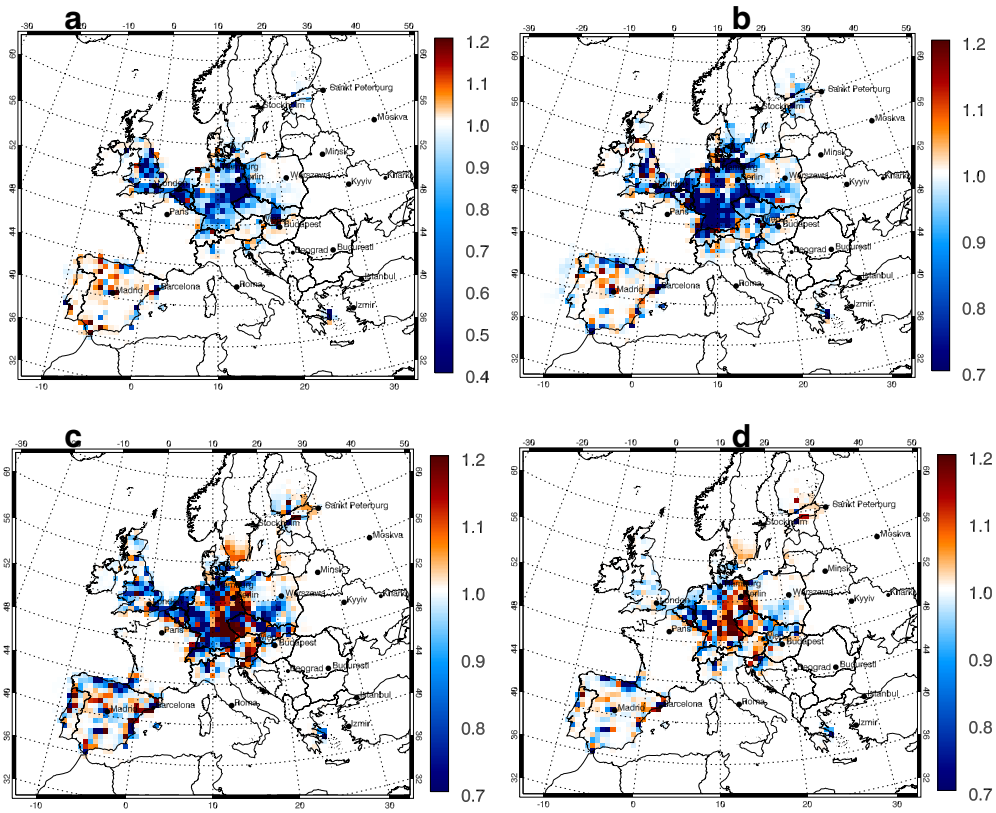

Fig. 10. Emission correction factors for **(a)** sulfur dioxide, **(b)** nitrogen dioxide, **(c)** terminal alkenes, and **(d)** isoprene at the surface layer, analysed by joint initial value/emission rate optimisation with 24 h assimilation interval placed at 17 August 1997.

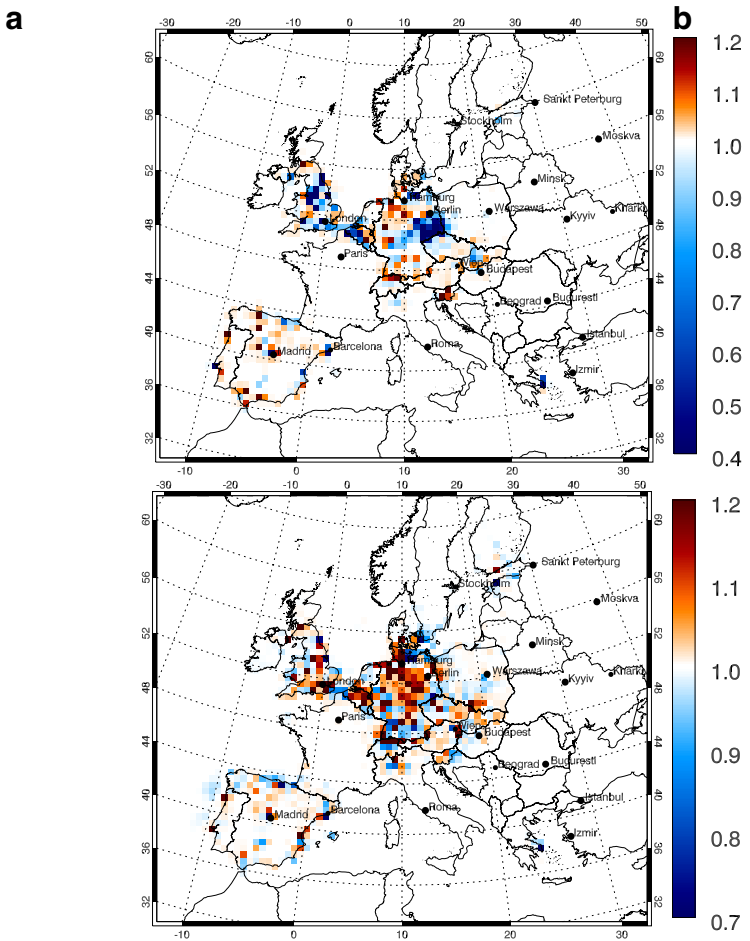

Fig. 11. Emission correction factors for **(a)** sulfur dioxide and **(b)** nitrogen dioxide as for Fig. 10, however with emission inventory reduced by 50% throughout the case study, for convergence control.

ACPD
7, 1725–1783, 2007

EGU

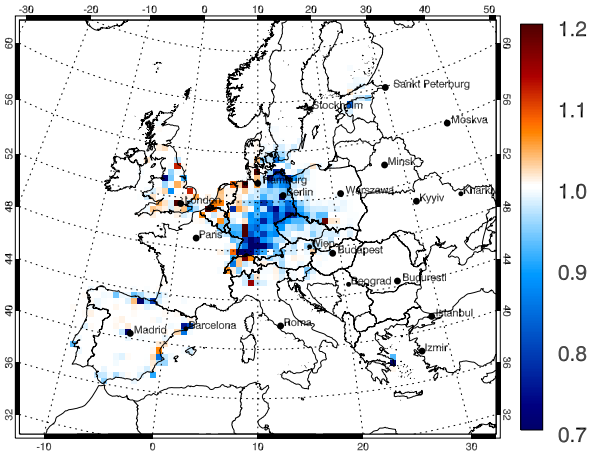

Fig. 12. Emission correction factors for nitrogen dioxide as for Fig. 10, however without any assimilation of NO_x observations throughout the case study.