

# Improving a complex chemistry-transport model by advanced optimization techniques

H. Elbern

*Institute for Geophysics and Meteorology, Project EURAD*

*University of Cologne, Cologne, FRG, he@eurad.uni-koeln.de*

**Keywords:** data assimilation, emission rate optimization, inversion, adjoint modelling

## 1 Introduction

Chemistry-transport models (CTMs) have, as many other models, manifold sources of uncertainties and errors. The model skill is not only dependent on completeness of the model formulation and grid resolution, but also on various model parameters, which take their part in controlling the simulation. The focus of the paper is placed on the variational calculus as a methodology to improve the skill and reliability of complex CTMs. With a systematic application of this technique, measurements are used to systematically (1) identify model deficiencies and uncertainties, and (2) optimize misspecified parameters to assure a better simulation quality and analysis of chemical states.

While data assimilation is widely associated with identification of initial values or state analysis, space-time assimilation algorithms can also be taken for the optimization of other modeled parameters, like emission rates, deposition velocities, boundary values and others. The method adopted to solve this inversion problem is a first real world implementation of the four-dimensional variational data assimilation algorithm (4D-var) based on the various adjoint model operators, which modify the modeled constituents.

The variational calculus has first been applied to atmospheric chemistry modelling by Fisher and Lary (1995) with a reduced stratospheric mechanism. Tropospheric applications have been described in Elbern et al. (1997) for a tropospheric box model, and in Elbern and Schmidt (1999, 2001) for the three-dimensional University of Cologne Eulerian CTM EURAD. All these studies focus on initial value or chemical state analysis. In polluted areas, however, emission rates can be of higher importance on a time scale larger than half a day (Elbern and Schmidt, 2001). Emission rate estimation with advanced modelling techniques is an active field of research since the work of Marchuk (1974), who introduced the variational calculus for the solution of the problem. Later studies in that field are due to Newsam and Enting (1988), Enting and Newsam (1990), Bousquet et al. (1999a,b), Gloor et al. (1999), Kaminski et al. (1999a,b), and Zhang and Heemink (1999). All these studies have in common to optimize the emission rates of the observed species. Van Loon et al. (2000) and Elbern et al. (2000) demonstrated for the first time, that, with suitably chosen regularization rules, emission rates of precursor species of ozone can be estimated by ozone observations. In the former case a reduced rank Kalman filter is implemented, while in the latter study a four dimensional variational approach is adopted. In this paper results will be presented for the joint optimization of initial values and emission rates. Misspecified emission parameters are an important error source for tropospheric CTMs.

## 2 Theory

The variational data assimilation method is likewise applicable to the initial value problem and parameter optimizations, as for emission rates. In our case the emission rates are subject to optimization. This section presents a brief outline of the theoretical background. For a review of variational parameter optimization studies see for example Navon (1997). A distance function or objective function may be defined as follows:

$$J(x, e(t)) = \frac{1}{2} \int_0^N (e_b(t) - e(t))^T K^{-1} (e_b(t) - e(t)) dt + \frac{1}{2} \int_0^N (\tilde{x}(t) - x(t))^T R^{-1} (\tilde{x}(t) - x(t)) dt$$

where  $J$  is a scalar functional defined on the time interval  $0 \leq t \leq N$  dependent on the vector valued state variable  $x(t)$ , and a parameter  $e(t)$  to be optimized. The error covariance matrices of the first guess or background emission rates  $e_b(t)$  and observations are denoted  $K$  and  $R$ , respectively. The chemistry transport model (CTM) with inclusion of emissions is given by  $\frac{dx}{dt} = M(x) + e$ , where  $M$  acts as a generally nonlinear model operator and  $e$  is in our case the vector of emission rates. Both terms uniquely define the state variable  $x(t)$  at time  $t$ , after an ever fixed initial state  $x(0)$  is provided. A detailed description of the theoretical background and implementation may be found in Elbern et al. (2000).

### 3 Assimilation system components

Four components may be identified to outline the design of the variational chemistry data assimilation algorithm: (1) the forward model, (2) the adjoint of its tangent linear version, (3) the background error covariance matrix, and (4) the minimization routine. The EURAD CTM2 is a comprehensive tropospheric Eulerian model operating on the mesoscale-alpha (Chang et al., 1987; Hass, 1991). The chemistry transport model calculates the transport, diffusion, and gas phase transformation of about 60 chemical species with 158 reactions. The present grid configuration is  $77 \times 67 \times 15$  with 54 km grid resolution and 100 hPa model top. The associated adjoint operators include the gas phase mechanism, the transport schemes and an implicit vertical diffusion scheme.

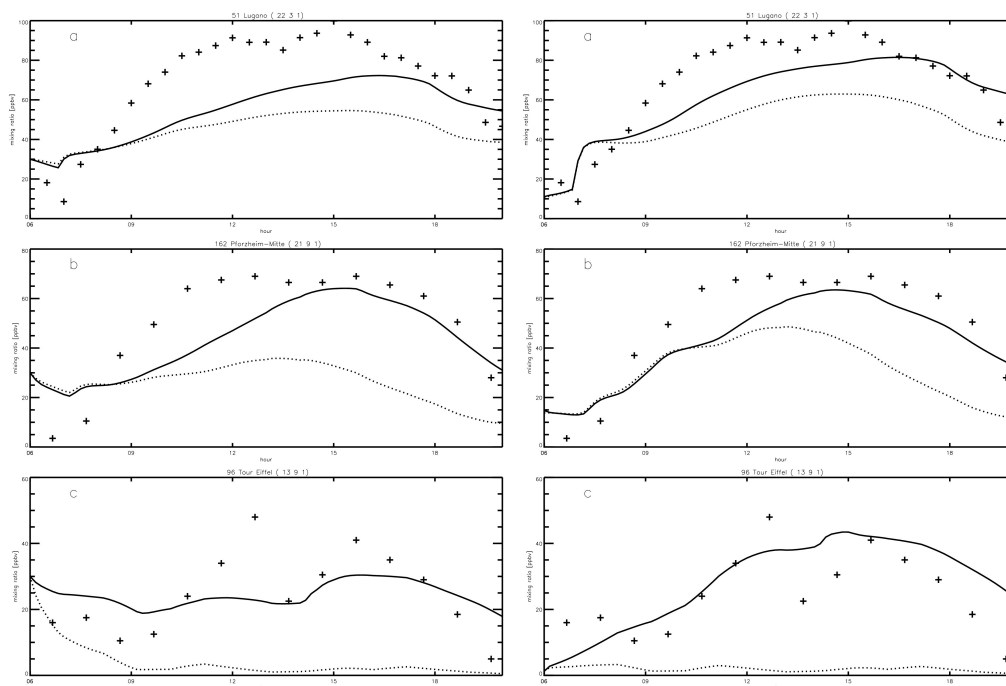
To ease computational burden of the assimilation process, most experiments are performed in a sub-domain with the same resolution. Another component is the emission module which includes 19 emitted species. A more detailed description may be found in Hass (1991). The emission data in this study are taken 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).

The development of the adjoint of the CTM2 comprises the coding and implementation of the adjoint operators of and given in(). The adjoint chemistry was coded by hand, while for the advection and diffusion routines the AMC adjoint model compiler (Giering and Kaminski, 1998) and Oyssee differentiation system (Faure and Papegay, 1998) were used for adjoint compilation. The correctness of the adjoint code was examined by the method proposed by Chao and Chang (1992). The gas phase chemistry solver and the implicit vertical diffusion operator apply adaptive time step techniques. During backward integration the same time steps are taken as determined by the forward integration. The quasi-newton limited memory L-BFGS algorithm devised in Nocedal (1980) and Liu and Nocedal (1989) is applied for the minimization.

### 4 Observations and meteorological situation

Observations of chemical constituents were obtained from various sources. A coarse coverage of Europe by observation sites is provided by EMEP data sets with hourly mean values of ozone measurements at 82 locations. In general EMEP stations are deployed to observe chemical background conditions with enhanced spatial representativeness of individual observation sites. On average, records of about 400 observation sites were available each day of the case study. Estimated accuracy is about 10 or at least 2 ppbv (parts per billion by volume) for ozone (personal communication V. Mohnen, Fraunhofer Institute for Environmental Research). An additional portion of the error is assumed to be due to the poor spatial representativeness as implied by a horizontal resolution of 54 km. This is a special problem for the use of urban and suburban measurement sites, although observations close to areas with elevated traffic load were omitted.

A long lasting episode of elevated ozone levels over central Europe, which took place from 3 to 20 August 1997, is selected for the application of the chemistry 4D-var system. The mesoscale meteorological simulations of this time span are made by MM5 which is restarted every 48 hours starting from August 1 0000 GMT to August 20 2400 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 August 3.



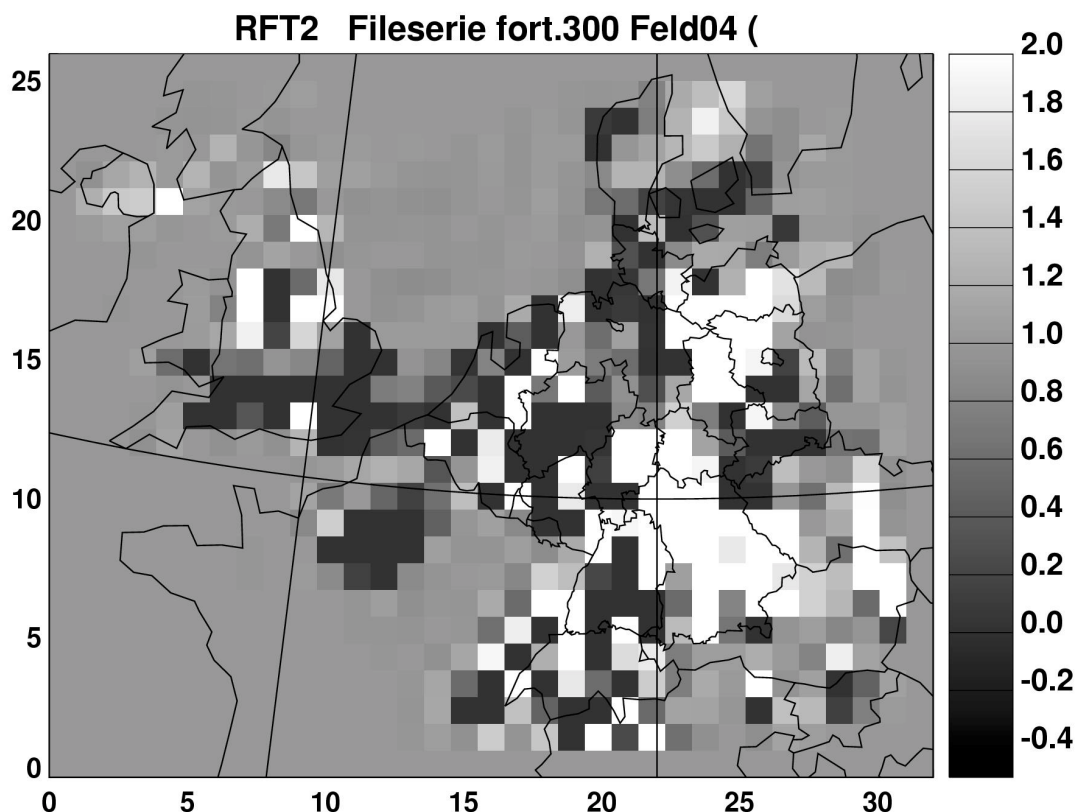
**Figure 1** Model simulations with emission optimization (solid lines) and without (dotted lines) for Lugano (top panels), Pforzheim (middle panels), and the lower platform of the Eiffel tour (bottom panels) for August 05, 06:00 to 20:00, 1977. Crosses denote ozone measurements. Left panel column depicts simulations without prior initial value optimization, right panel column with prior initial value optimization.

## 5 Results

Prior to the optimization runs a traditional forward simulation (reference run) was performed with four days spin-off time. Optimization runs were made in two different ways: the first procedure starts with initial values as obtained by the reference run. Any optimization is then achieved by modification of the emission rates. The assimilation window for emission rate optimization is selected to embrace the daytime traffic emissions, and starts at 06:00 UTC in the morning and ends at 20:00 UTC. Figure 1 exhibits results for stations with obviously strongly misspecified emission rates, which are Lugano, Pforzheim, and the lower platform of the Eiffel tour. The left panel column displays results from the pure emission optimization runs. The simulation performance is markedly improved in the course of the assimilation interval. Large initial value discrepancies between the reference run and observations during the first hours stem from the lagging effect of emission modifications.

The right column of Figure 1 shows the same case with prior initial value optimization, where the assimilation interval was restricted to 6 hours. It is seen that the model-observation discrepancies are reduced as compared to the former case. This results in an overall better performance within the entire 14 hours analysis interval. The case of the Eiffel tour station is of special interest, as the initially simulated NO<sub>x</sub> emissions are suppressing the ozone levels below 5 ppbv prior to the optimization procedure. Hence, the model simulation portrays a different chemical regime as monitored by the observations. Nevertheless, the variational emission optimization is able to drastically improve the model performance, as can be verified from the lower panels in Figure 1.

Figure 2 gives the amplification factors for NO emissions for the case of prior optimization of initial values. In most areas with occurring modifications, as given by the areas with available measurements stations, an augmentation of the reference emissions is analysed, especially in the southern and eastern part of Germany. The strongly overspecified NO emissions in the Paris area, as also revealed by the Eiffel tour station, is given by the darker grid cells, analysed in the Ile de France region. In summary it can be stated, that the emission rate optimization by the 4D-var calculus is able to optimize emission rates of ozone precursor species, even if only ozone as the product species is observed.



**Figure 2** Optimized emission factors of NO after assimilation with prior initial value optimization. Factor 1 shading either indicates proper emission rates or unobserved regions.

## References

- Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray, Inverse modeling of annual CO<sub>2</sub> sources and sinks, 1. Method and control inversion, *J. Geophys. Res.*, 104(D21), 26161-26178, 1999a.
- Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray, Inverse modeling of annual CO<sub>2</sub> sources and sinks, 2. Sensitivity study, *J. Geophys. Res.*, 104(D21), 26179-26193, 1999b.
- Chang, J.S., R.A. Brost, I.S.A. Isaksen, S. Madronich, P. Middleton, W.R. Stockwell, and C.J. Walcek, A three-dimensional Eulerian acid deposition model: physical concepts and formulation, *J. Geophys. Res.*, 92, 14,618--14,700, 1987.
- Chao, W.C. and L.-P. Chang, Development of a four-dimensional variational analysis system using the adjoint method at GLA. Part 1: Dynamics. *Mon. Wea. Rev.*, 120, 1661-1673, 1992.
- Elbern, H., and H. Schmidt, Ozone episode analysis by four--dimensional variational chemistry data assimilation, *J. Geophys. Res.*, 106, 3569-3590, 2001.
- Elbern, H., H. Schmidt, O. Talagrand, A. Ebel, 4D--variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Mod. Softw.*, 15, 539-548, 2000.
- Elbern, H., and H. Schmidt, A four--dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.*, 104, 18 583--18 598, 1999.
- Elbern H., H. Schmidt, and A. Ebel; Variational data assimilation for tropospheric chemistry modeling. *J. Geophys. Res.*, 102, D13, 15,967--15,985, 1997.
- Enting, I.G., and G.N. Newsam, Inverse problems in atmospheric constituent studies: II. Sources in the free atmosphere, *Inverse Problems*, 6, 349-362, 1990.
- Faure, C., and Y. Papegay, Oyssee Version 1.7, the user's reference manual, *Rapport Technique 0224*, INRIA, Sophia Antipolis, France, 1998.
- Fisher, M., and D.J. Lary, Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, 121, 1681-1704, 1995.
- Giering, R., and T. Kaminski, Recipes for adjoint code construction, *ACM Trans. Math. Software*, 24, 437--474, 1998.

- Gloor, M., Fan, S.-M., S. Pacala, J. Sarmiento, and M. Ramonet, 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 CO<sub>2</sub> on a continental scale, *J. Geophys. Res.*, 104(D12), 14245-14260, 1999.
- Hass, H., Description of the EURAD Chemistry--Transport--Model Version 2 (CTM2), *Mitteilungen aus dem Institut für Geophysik und Meteorologie der Universität zu Köln*, Nr. 83, edited by A. Ebel, F.M. Neubauer, P. Speth, 100 pp, 1991.
- Kaminski, Th., M. Heimann, R. Giering, A coarse grid three-dimensional global inverse model of the atmospheric transport, Inverse problems in atmospheric constituent studies: 1. Adjoint model and Jacobian matrix, *J. Geophys. Res.*, 104(D5), 18535-18553, 1999a.
- Kaminski, Th., M. Heimann, R. Giering, A coarse grid three-dimensional global inverse model of the atmospheric transport, 2. Inversion of the transport of CO<sub>2</sub> in the 1980s, *J. Geophys. Res.*, 104(D5), 18555-18581, 1999b.
- Liu, D.C., and J. Nocedal, On the limited memory BFGS method for large scale minimization. *Math. Prog.*, 45, 503--528, 1989.
- Marchuk, G., Numerical solution of problems of dynamics of the ocean and the atmosphere. *Gidrometeoizadat*, Leningrad (in Russian), 1974.
- Memmesheimer, M., H. Hass, J. Tippke, A. Ebel, Modeling of episodic emission data for Europe with the EURAD Emission Model (EEM), in *Regional Photochemical Measurement and Modeling Studies*. Vol. 2, edited by A. J. Ranzieri, P.A. Solomon, Air Waste Management Association, Pittsburgh, USA, 495--499, 1995.
- Navon, I.M., practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dynam. Atmos. Oceans*, 27, 55-79, 1997.
- Newsam, G.N., and I.G. Enting, Inverse problems in atmospheric constituent studies: I. Determination of surface sources under a diffusive transport approximation, *Inverse Problems*, 4, 1037-1054, 1988.
- Nocedal, J., Updating quasi-Newton matrices with limited storage. *Mathem. Comp.*, 35, 773-782, 1980.
- van Loon, M., P.J.H. Builtjes, A.J. Segers, Data assimilation applied to LOTOS: first experiences, *Environ. Mod. Softw.*, 15, 603-609, 2000.
- Zhang, X.F., and A.W. Heemink, The application of Kalman filter in the evaluation of the CH<sub>4</sub> budget in Europe. In: *Air Pollution IV, monitoring, simulation and control*, 155-164, Wessex Institute of Technology, 1999.