Inverse Modelling of Atmospheric Trace Substances on the Regional Scale with Lagrangian Models

A contribution to subprojects GLOREAM and GENEMIS-2

P. Seibert, A. Frank and H. Kromp-Kolb

Institute of Meteorology and Physics, University of Agricultural Sciences, Vienna, Austria
petra.seibert@boku.ac.at

Introduction

Measurements can not only be used as stand-alone data or for model evaluation. If we trust the measured data and our models sufficiently well, we can use them to infer source properties. The method to do this is called inverse modelling; it is an optimisation technique. Sources (or source parameters) are adjusted so that the differences between model output and measurements are minimised. Inverse problems can be ill-conditioned or (partly) underdetermined and it may be necessary to add a-priori knowledge to constrain the solution. Inverse modelling is closely related to data assimilation, especially variational and Bayesian techniques where the control variables to be adjusted are concentration fields rather than sources.

Regular air pollution simulations use meteorological fields and emission inventories as input, and calculate time-dependent concentration fields as output, using suitable models for the relevant processes such as transport, diffusion, deposition and chemical transformation. Observations of atmospheric trace substances can then be compared with model output, and conclusions on model performance can be drawn. In a qualitative way, possible deficits of the model or its input can be identified. Inverse modelling aims at using observations in a quantitative way, in order to extract quantitative information on emissions (or, in principle, model parameters or other model input such as meteorological variables).

Inverse modelling is an optimisation technique: the control variables (e.g., the emissions) are varied to find values leading to the minimisation of a suitable cost function. The main contribution to the cost function is a measure of the deviation between modelled and observed concentrations. Deviation of the control variables from some a-priori expectation (e.g., a first guess) can be another contribution to the cost function.

The optimisation problem can be solved in different ways. The most general method is the determination of the gradient of the cost function with respect to the control variables and iterative minimisation. The method we selected computes the full source-receptor matrix and minimises the cost function analytically, solving a linear system of equations.

Objectives

This contribution aims at the development of inverse modelling methods to derive information on emissions from measurements in the regional scale. Such methods shall be applied to suitable data sets.

P.M. Midgley, M. Reuther (Eds.)
Proceedings from the EUROTRAC-2 Symposium 2002

Activities and Results

General approach: calculation of source-receptor matrix and regularised inversion

There are many possible approaches to inverse modelling (see, e.g., Seibert, 2001b). The one we have chosen is based on the assumption of a linear source-receptor relationship. This condition is appropriate for all substances in the atmosphere which do not undergo nonlinear chemical transformations, especially radionuclides, typical artificial tracers, many aerosol particles and slowly reacting species such as many greenhouse gases or CFCs. More reactive substances can still be treated if their transformation rates can be prescribed, e.g., from climatological OH concentrations. Even if the linearity is not so well fulfilled, as in the case of sulphur, useful results can still be obtained as shown below.

The linearity condition allows to split the problem in two independent steps. The first step is the determination of a source-receptor matrix (SRM) whose elements give the sensitivity of the receptor elements (e.g., measured concentrations) to the source elements (e.g., the emissions from each grid cell). The second step solves the linear problem of finding the unknown source elements which lead to those concentrations which best fit the observations.

The SRM calculation can be based on any kind of dispersion model in either forward or backward (adjoint) mode. The computational burden is proportional to the number of source elements in the former and the number of receptor elements in the latter case. Furthermore, using a Lagrangian (particle or simple trajectory) model, the backward, receptor-oriented mode is also the most accurate one for point measurements in combination with gridded sources. Details are presented below in the context of applications.

SRMs based on simple back trajectories and application to sulphur in Europe

Simple back trajectories, without dispersion calculations, have long been used for the interpretation of chemical monitoring data. Potential source areas have been calculated using so-called trajectory statistics (Seibert et al., 1994; Stohl, 1996). Often, long-term archives of trajectory data exist. Therefore it is of interest how this information content can be used in the best way, though redoing calculations with a full dispersion model would be much more accurate. As shown in Seibert (1999), they can be used to calculate elements of a source-receptor matrix:

$$m_{ijl} = \frac{\partial y_l}{\partial x_{ij}} = \Delta\tau_{ijl}\; e^{-k_1 \tau_{ijl}}\; (1 - e^{-k_2 \tau_{ijl}})$$

where m_{ijl} is an elements of the SRM (ij denotes the source location and l the observation), y denotes the observed concentration normalised by an averaged mixing height (in units of mass per volume [STP]), x denotes the source (in units of mass per volume [STP] and time), $\Delta\tau$ is the residence time of the trajectories belonging to observation l in source area ij, and τ is the travel time from the source area to the receptor. The exponential terms are a climatological representation of the residence time of the substance under consideration (k_1 is the respective 1/e-time) and of the formation of the substance (here: SO_4) from a primary pollutant (here: SO_2), with a rate constant of k_2.

In the second step, the inversion is performed with respect to the vector $\boldsymbol{x}$ of sources, using the observations in vector $\boldsymbol{y}^o$. In order to guarantee a stable solution despite the fact that the source-receptor matrix $\boldsymbol{M}$ is ill-conditioned, a regularisation with two additional terms in the cost function is applied. This means that the minimum of the cost function J defined as follows is to be found:

$$J \equiv \underbrace{\|\boldsymbol{M}\boldsymbol{x} - \boldsymbol{y}^o\|^2}_{\text{RMSE}} + \underbrace{\theta_1^2 \|\boldsymbol{x}\|^2}_{\text{variance of } \boldsymbol{x}} + \underbrace{\theta_2^2 \, \|\boldsymbol{D}\boldsymbol{x}\|^2}_{\text{roughness } \nabla^2 \boldsymbol{x}} \stackrel{!}{=} \text{Min}$$

The first contribution to *J* gives the RMSE; the second and third term are for regularisation. The first regularisation term minimises the total variance of the solution vector (Tikhonov regularisation), while the second regularisation term minimises its roughness (using a matrix operator ***D***). This leads to the final linear system of equations to be solved:

$$(\boldsymbol{M}^T\boldsymbol{M} + \theta_1^2 \; \boldsymbol{I} + \theta_2^2 \; \boldsymbol{D}^T\boldsymbol{D}) \; \boldsymbol{x} = \boldsymbol{M}^T\boldsymbol{y}^o$$

where ***I*** is the unit matrix. The weights $\theta^2_{1,2}$ are presently determined by trial and error.

This procedure was applied to EMEP's daily measurements of airborne sulphate at 13 stations in Europe in combination with 4-day 3D backward trajectories to infer the annual mean sulphur emission pattern over Europe. The main centres of sulphur emission in Europe are recovered (except the far Southeast, where there were insufficient data), but – probably due to the regularisation – the variance of the source distribution is underestimated (Figure 1). Errors in the emission field are due to errors in the SRM and the measurements. A test with perfect data showed that emissions could be reconstructed perfectly with this method, whereas simple trajectory statistics would explain only 25% of the variance.

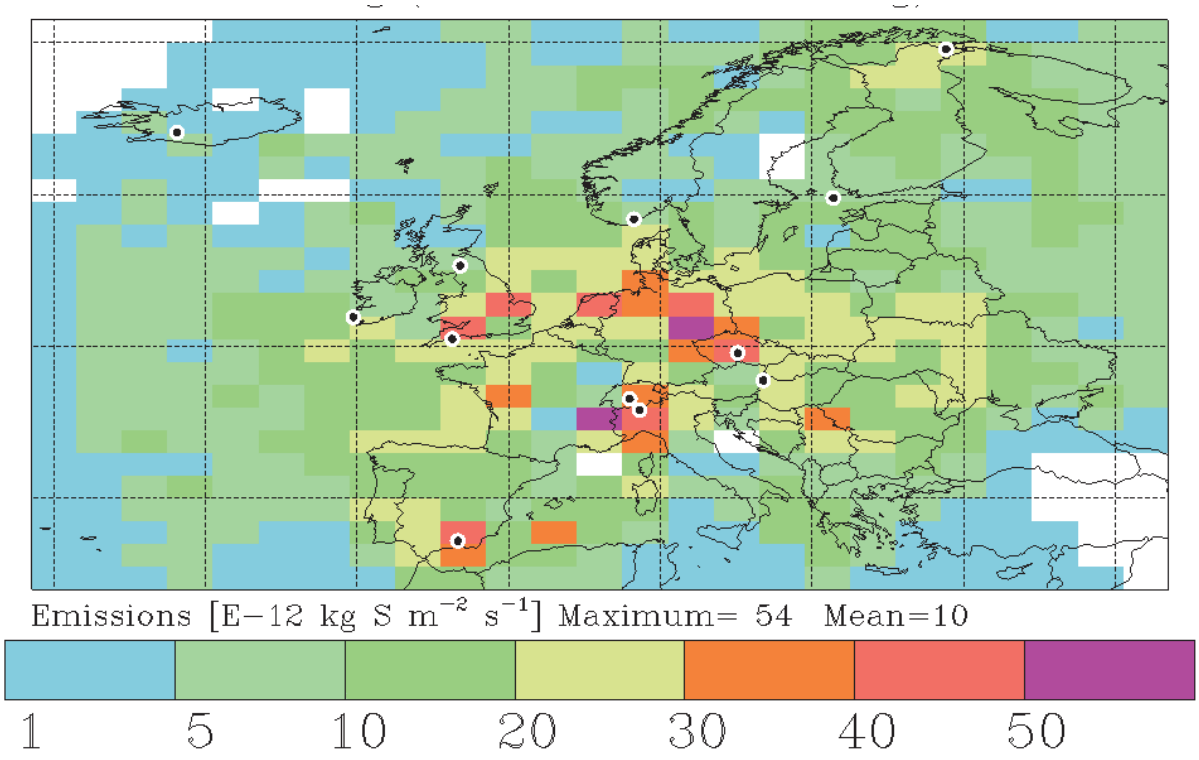

Figure 1. Distribution of annual sulphur emissions (in 10^{-12} kg S $m^{-2}s^{-1}$ in Europe as derived with the inverse modelling method from simple back trajectories and daily sulphate measurements at 13 EMEP stations during one year (adapted from Seibert, 1999).

SRMs based on the FLEXPART dispersion model and application to ETEX

In the first of the two releases in the European Tracer Experiment ETEX, an inert tracer was emitted near ground at a site in NW France during 16 hours and its concentration was sampled by 168 station in large areas of Europe in the subsequent days, with a temporal resolution of 3 hours. An inverse modelling was carried out for this release, using the Lagrangian particle dispersion model FLEXPART (Stohl et al., 1998) in backward mode to calculate the source-

receptor matrix. For this purpose, a pseudotracer was emitted from each measurement site during the measurement interval, transported backward with the full dispersion model (including turbulence – deposition and decay were not relevant here though they are implemented) and sampled on a 1-degree grid with 1 hour resolution. It was shown that pseudotracer concentrations in grid cells represent the SRM values after some simple scaling (Seibert and Stohl, 2000; Seibert, 2001), which are basically the respective particle residence times (similar to the simple trajectory case). The extension to the case with loss processes has also been derived and will be published soon.

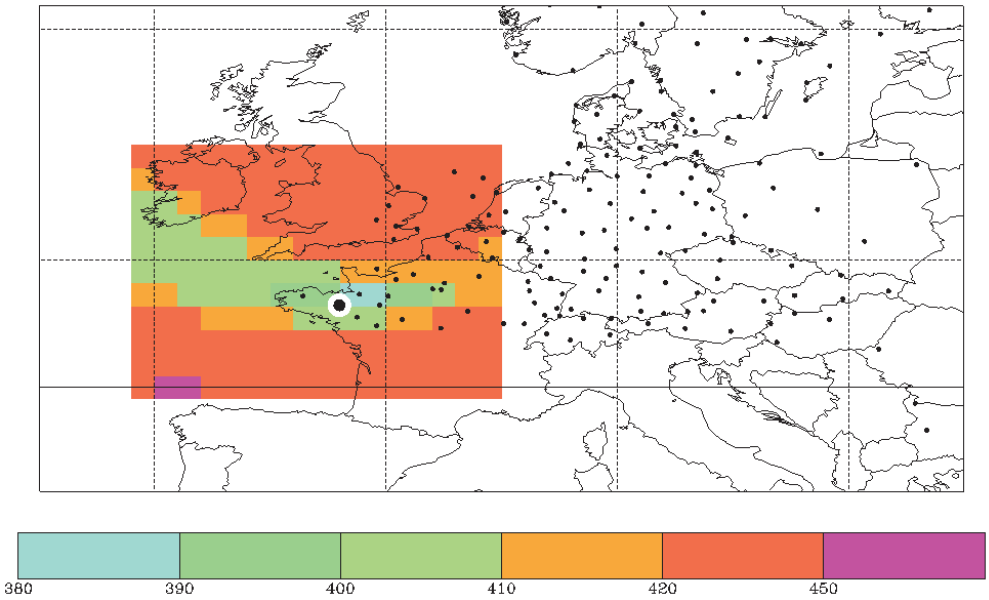

Figure 2. Distribution of the RMSE (in ngm^{-2}) between observed tracer concentrations and modelled ones, resulting from inverse modelling of the temporal evolution of the ETEX first release source term if source is the respective grid cell. The big dot is the true release site, the small dots are the measurement sites. (From Seibert and Frank, 2001b)

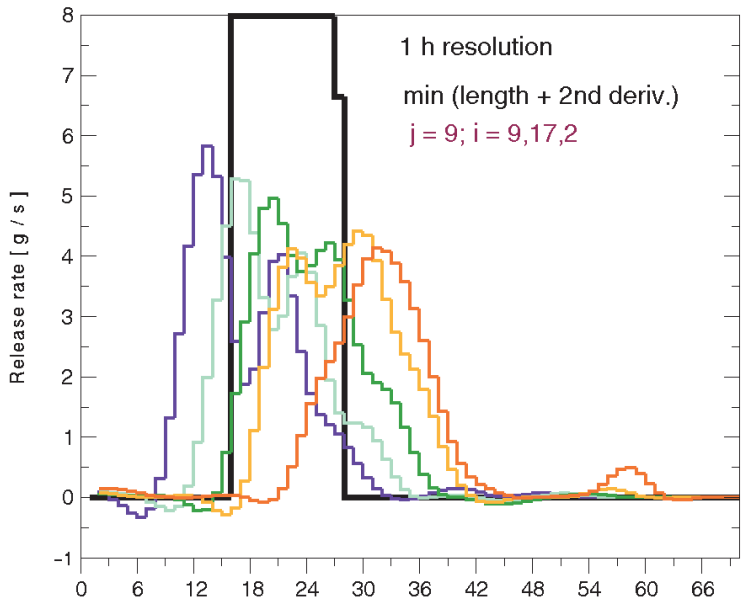

Figure 3. Temporal evolution of reconstructed tracer emissions during the ETEX first release for every other releases assumed in different cells of the grid row containing the real source location. Index i=1 corresponds to the westernmost grid cell of the coloured area in Figure 2 and is represented by the blue curve. Further curves are shifted by 2 grid distances each. The black line represents the true release rate.

Using the same inversion technique as for the sulphur study, it is possible to reconstruct the temporal evolution of the source with its location given, or the geographical distribution when the temporal shape is given. In order to reconstruct the temporal shape and the location while using the a-priori knowledge that the source is a point source (or in our discretisation, a one grid-cell source), the temporal shape is reconstructed for all possible grid cells, and the resulting RMSE and correlation coefficients between observed concentrations and concentrations resulting from the application of the SRM to the reconstructed source function are calculated (see Figure 2); the grid cell with the smallest RMSE (largest correlation) would then be the reconstructed source location. We can see that the method works well. The location is found within about 1 grid distance from the true one, and the temporal shape (Figure 3) is found with reasonable accuracy. There is a certain underestimation of the total released mass, likely related again to the regularisation. The fuzzy end is related to late observations of the tracer which are incompatible with forward modelling results.

Applications in the context of the Comprehensive Nuclear Test Ban Treaty verification

The Comprehensive Nuclear Test Ban Treaty (CTBT) relies on four monitoring techniques for the detection of nuclear explosions: seismic and hydroacoustic waves, infrasound, and radionuclides. A global network of 79 monitoring stations for airborne radionuclides is being set up; stations will take 24-hour samples. Of course, the question arises where a possible source may have been located if suspicious nuclides are found in samples. To answer this question, atmospheric transport models are employed to calculate so-called 'Fields of Regard' at the International Data Centre of the CTBT Provisional Technical Secretariat (PTS) in Vienna, where all monitoring data are collected in real time. These fields are basically potential source areas (volumes) for a given sample. The natural way to calculate these fields are backward tracer transport simulations in the mode which we developed for the ETEX test case. Recently, the PTS conducted an evaluation of their models with the help of external experts running their own models (CTBT, 2001). We participated with the model FLEXPART in this expert group. One of the interesting results was that for this type of application, extreme dilution ratios are of interest. However, atmospheric modellers have yet put no effort in the determination of the reliability of their models for such situations. We found that for a Lagrangian particle model, the number of particles required for a statistically stable situation can be quite high, on the order of hundreds of thousands, but depending strongly on the exact specifications (Seibert and Frank, 2001a).

Convection scheme for FLEXPART

Vertical transports by convective clouds cause nonlocal mixing in the atmosphere, which cannot be represented by the usual turbulent mixing parameterisations in Eulerian as well as Lagrangian models. Atmospheric transport modelling in the tropical belt as well as long-term simulations and simulations which focus on vertical transports in the free atmosphere require thus the implementation of a convection scheme. Therefore, global, climate-oriented models typically include such parameterisations. On the background of the EU Framework project STACCATO and also our activities in the context of the CTBT verification, a convection scheme was developed for FLEXPART (Seibert et al., 2002). It relies on the Emanuel scheme (Emanuel and Zivkovic-Rothman, 1999) to calculate a redistribution matrix. Particles in convective grid columns are redistributed stochastically according to probabilities given by this matrix in forward mode and according to the transposed matrix in backward mode.

Conclusions

A new method for inverse modelling of atmospheric trace substances has been developed. It is based on the calculation of a source-receptor matrix (SRM) and subsequent analytical minimisation of a cost function made up of the squared observation-model errors and suitable

regularisation terms representing additional a priori information. The SRM can be approximated on the basis of simple back trajectories, a procedure which offers an attractive alternative to trajectory statistics. Otherwise, it was calculated with a Lagrangian particle model (forward or backward mode possible). The method has so far been applied to sulphur in Europe (with simple trajectories) and the ETEX first release (Lagrangian particle model). SRM calculations have also been performed in the context of the verification of the Comprehensive Nuclear Test Ban Treaty, where the high accuracy of this approach is very welcome. More applications, both of SRM calculation and of subsequent inversion of sources, are pending.

Acknowledgements

This work was funded by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (P12595-GEO), the Austrian Federal Ministry of Education, Science and Culture (GZ 76.017/1-VIII/A/5) and the European Commission (project STACCATO, EVK2-CT-1999-00050).

References

CTBTO, 2001: Evaluation of the atmospheric transport modelling tools used at the Provisional Technical Secretariat. Report to the PTS of an Ad Hoc Expert Group on the Evaluation of Atmospheric Models Used at the PTS for Radionuclide Transport Analyses. Preparatory Commission for the ComprehensiveNuclear-Test Ban Treaty Organization, Vienna, Report CTBT/EVA/EG-1/1, 73 pp.

Emanuel, Kerry A., and M. Zivkovic-Rothman, 1999: Development and evaluation of a convection scheme for use in climate models. *J. Atmos. Sci.*, **56**:1766-1782.

Seibert, P., 1999: Inverse modelling of sulfur emissions in Europe based on trajectories. In P. Kasibhatla, M. Heimann, P. Rayner, N. Mahowald, R . G. Prinn, and D. E. Hartley, editors, *Inverse Methods in Global Biogeochemical Cycles*, AGU Geophysical Monograph Vol. 114, ISBN 0-87590-097-6, Washington., 147-154.

Seibert, P., 2001a: Inverse modelling with a Lagrangian particle dispersion model: application to point releases over limited time intervals. In: F.A. Schiermeier and S.-E. Gryning, editors, *Air Pollution Modeling and its Application XIV*, Kluwer Academic Publ., 381-390.

Seibert, P., 2001b: Methods for source determination in the context of the CTBT radionuclide monitoring system**.** In *Proceedings Informal Workshop on Meteorological Modelling in Support of CTBT Verification, Vienna, December 2000*. On CDROM and on-line.

Seibert, P., H. Kromp-Kolb, U. Baltensperger, D.T. Jost, M. Schwikowski, A. Kasper, and H. Puxbaum, 1994: Trajectory analysis of aerosol measurements at high alpine sites. In *Proc. EUROTRAC Symposium* 1994, 689-693.

Seibert, P., and A. Stohl, 2000: Inverse modelling of the ETEX-1 release with a Lagrangian particle model**.** In G. Barone, P.J. Builtjes, and G. Giunta, editors, *Proceedings 3rd GLOREAM Workshop*, Sept. 1999, Ischia , Italy, 95-105.

Seibert, P., and A. Frank, 2001a: Backward Atmospheric Transport and Dispersion Modelling with FLEXPART in the Context of Evaluation of Atmospheric Transport Modelling Tools at CTBTO/PTS. Annex to *Evaluation of Atmospheric Transport Modelling Tools Used at the Provisional Technical Secretariat*, Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization, Paper EVAL-102/01/450, 8 p..

Seibert, P., and A. Frank, 2001b: Backward Lagrangian particle modelling and source-receptor matrices as a tool in environmental meteorology. Proc. D-A-CH Meteorologentagung, Wien, *Österreichische Beiträge zu Meteorologie und Geophysik*, **27** (on CDROM).

Seibert, P., B.C. Krüger and A. Frank, 2002: Parameterisation of convective mixing in a Lagrangian particle dispersion model. *Proceedings. 5th GLOREAM Workshop, Wengen (CH)*, on CDROM and on-line.

Stohl, A. 1996: Trajectory statistics - a new method to establish source-receptor relationships of air pollutants and its application to the transport of particulate sulfate in Europe. *Atmos. Environ.*, **30**, 579-587.

Stohl A., Markus Hittenberger, and Gerhard Wotawa, 1988: Validation of the Lagrangian particle dispersion model Flexpart against large-scale tracer experiment data. *Atmos. Environ.*, **32**, 4245-4264.