Methods for source determination in the context of the CTBT radionuclide monitoring system

Petra Seibert

Institute of Meteorology und Physics, University of Agricultural Sciences Vienna
Türkenschanzstr. 18, A-1180 Wien, Austria. E-Mail: petra.seibert@boku.ac.at

1. Introduction

A variety of methods applicable to the problem of source determination or inverse modelling of atmospheric trace constituents has been suggested and used so far. They include trajectory statistics (Stohl, 1998), inverse methods based on the source-receptor matrix, and iterative methods based on adjoint Eulerian dispersion models. Source-receptor matrices (SRMs) can be computed by different kinds of models, in forward as well as in backward mode. The different methods are characterised and their advantages and drawbacks will be discussed, with special emphasis on CTBT verification.

2. The nature of the source determination problem

The source determination problem is sometimes associated with terms such as 'backward modelling' or 'backtracking'. Though dispersion models may be run in a backward mode, this is neither necessary nor sufficient for source determination. There is no direct way of calculating sources from observed concentrations, comparable to the forward simulations of concentrations fields from given sources. The source determination is an optimisation problem. We want to find the sources x leading to calculated concentrations y that fit best the observed concentrations y^o when being plugged in a transport & dispersion model. In mathematical terms, we want to minimise a cost function $J[\mathcal{M}(x), y^o]$ that measures the misfit of the model output $\mathcal{M}(x)$, using the source strength x (in general, a function of time and space) as the control variable. Eventually, we want to include further terms in the cost function to represent additional *a priori* knowledge or assumptions. The methods that can be used to find this minimum depend on whether the gradient of the cost function $\nabla_x J$ and the model operator $\mathcal{M}$ are linear with respect to x or not. In the nonlinear case, iterations are necessary, whereas in the linear case an analytical solution can be set up. The transport, dispersion, deposition and radioactive decay of chemically inert species (or species whose chemical transformations follow predescribed rates) is a linear problem, and thus so is the CTBT verification problem. The typical assumption for the cost function, namely to be equal to the sum of the squared deviations between model output and observations, also fulfils the linearity requirement. Replacing $\mathcal{M}$ with the SRM $\boldsymbol{M}$, and considering a regularisation term expressed as a matrix $\boldsymbol{D}$ applied to the source vector x, with regularisation parameter (weight) θ^2, this can be written as a formula:

$$J = (\boldsymbol{M}x - \boldsymbol{y}^o)^T(\boldsymbol{M}x - \boldsymbol{y}^o) + \theta^2(\boldsymbol{D}x)^T(\boldsymbol{D}x) = \text{Minimum!}$$

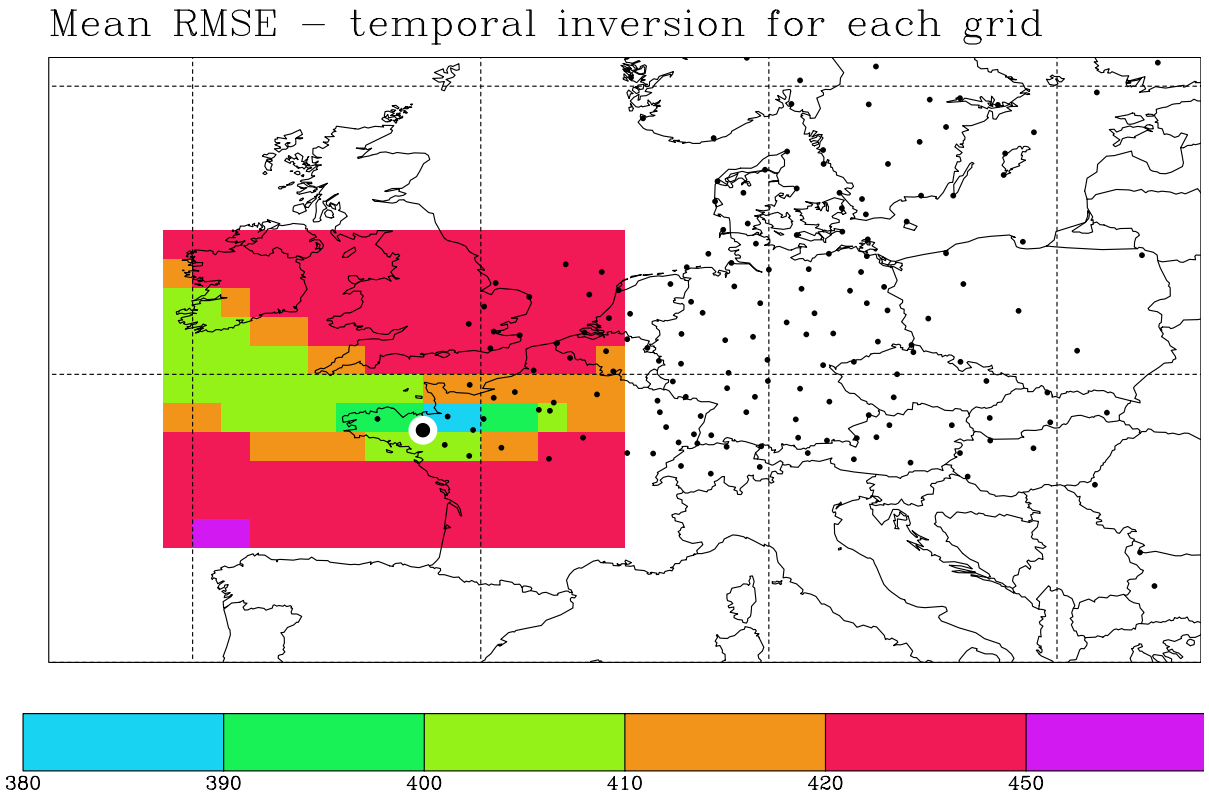
Of course, Bayesian methods can also be used for the optimisation (Høst, 1996), but it should be kept in mind that typical features of Bayesian methods can be incorporated also in the conceptionally more simple variational approach with generalised matrix inversion. For example, the cost function can be calculated to measure also the deviation to a first-guess solution, and error-correlation matrices can be introduced as weights (Menke, 1984).

3. Methods for finding the optimum

3.1 Iterative method

The iterative method starts from a first guess, calculates the gradient $\nabla_x J$ of the cost function with respect to the control variable(s), and then uses a suitable descent algorithm to go along the gradient in iterations until a minimum is found. In complex, nonlinear models the usual tool for the computation of the gradient is an adjoint version of the model, which runs backward. The model is expanded to include

Figure 1: *Mean error observerd vs. modelled concentrations, if the source is assumed to be in the respective grid cell, and its temporal evolution determined by regularised inversion. The small dots indicate the measurement sites, the big dot is the real source location. The units are pg m^{-3}.*



the cost function calculation, allowing an arbitrary, numerical formulation of this function. Of course, this method can be used also in a linear case (Robertson and Langner, 1998).

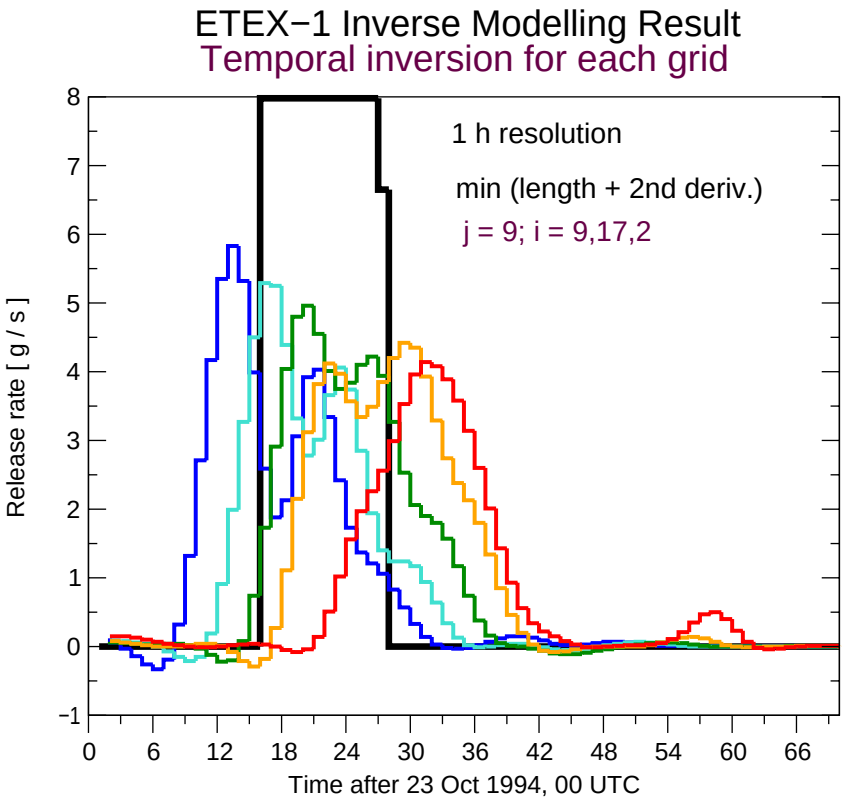
3.2 Source-receptor method

This method is typically used in cases with a linear source-receptor relationship. In this case, the complex model operator becomes a simple (but nevertheless potentially large) SRM $\boldsymbol{M}$, with the number of columns equal to the number of source elements and the number of rows equal to the number of receptor elements (measurements). The potential source in general is a four-dimensional field in time and space, that has to be divided into discrete cells. The measurement vector will typically be comprised of a number of stations and a number of measurement intervals at each station.

The SRM can be calculated by a variety of methods: Eulerian as well as Lagrangian dispersion models can be used, and they can be used in forward mode or in backward mode. The backward mode requires an adjoint model version, which is, however, not very different and relatively easy to formulate for a simple, linear problem. Once the SRM has been found, it can be inverted with a least-square algorithm. Often, especially if the source vector dimension is not small compared to the observation vector dimension, a generalised inversion with a regularisation term is necessary (Seibert, 1999, 2000). The *a priori* knowledge that the source is a point source is not so easily integrated in a regularised inversion as it is a nonlinear constraint. A possibility is to compute the SRM (possibly in backward mode) and then do forward calculations with this SRM and assumed source locations to dermine the respective cost function values J; the location yielding the smallest value of J is then easily found. In Seibert (2000), this approach was combined with an inversion for the temporal source variation only.

I would like to illustrate the SRM / regularised inversion method with an application to the first release of the European Tracer Experiment ETEX (Nodop et al., 1998). The SRM has been calculated with the

Figure 2: *Temporal evolution of the source as obtained from a regularised inversion (minimising the variance and the second derivative of the source), for several locations indicated by the grid number in x-direction as given in Figure 1, at the latitude where the RMSE is minimised. The green curve is the one that gives also the lowest RMSE. The real source shape is indicated by the bold black line.*



Lagrangian particle dispersion model FLEXPART (Stohl et al., 1998), running in backward mode. To make use of the knowledge that the source was a point source, each grid element was tried out as the location of the source (which means that the source is assumed to be spread out over one grid cell, in this case $1^\circ \times 1^\circ$, and its temporal evolution was determined by the inversion. For more details, see Seibert and Stohl (2000); Seibert (2000). Figure 1 shows the RMSE between observed tracer concentrations and modelled ones, with the reconstructed source term, for each source grid. There is a clear region where the source wouldbe relatively well compatible with the observations, and the grid cell with the smallest error is very close to the real source location. The temporal the resulting temporal distribution of the source for a few such source locations is shown in Figure 2, and also quite realistic.

3.3 Discussion of methods

Table 1 lists the characteristics of different source determination methods, along with authors who have applied them.

The iterative method is mandatory for nonlinear problems. In linear problems it can be used to avoid the matrix inversion, but this is at the cost of not really knowing whether a global minimum has been found. It may save disk space and – depending on the nature of the problem – also calculation time, but it gives also less insight. If another cost function is to be tried, the whole calculations have to be repeated.

The SRM method, on the other hand, gives an exact solution and also the possibility to check the stability of the solution analytically. It is easy to study the impact of each single observation. Changing the cost function, e. g., reassigning weights, is relatively inexpensive in computing time. This allows also to incorporate a non-linear cost function through iterations.

In the forward approach to SRM calculation, at each source a unit amount of tracer is released and sampled at the receptors. A separate species must be used (or a separate run performed) for each potential source. The computational effort is therefore determined primarily by the number of sources. In the backward or adjoint approach, tracer is released at each receptor, and a separate species (or run) is needed for each measurement. Thus, the preferred method depends on the ratio of the number of measurements to the number of potential sources.

Table 1: *Possible approaches to inverse modelling of atmospheric trace substances.*
FWD / BWD ... forward / backward
SRM ... source–receptor matrix, solve linear system of equations analytically
ITER ... iterative optimisation of a linear or nonlinear system (scalar cost function)

model type	**approach**		**# of runs (variables)**	**code required**	**type of problem**	**examples**
Eulerian	FWD	SRM	# sources	forward	linear	Mulholland & Seinfeld, AE 1995, 497
Eulerian	BWD	SRM	# receptors	fwd adjoint	linear	Pudykiewicz, AE 1998, 3039
Eulerian	BWD	ITER	2x # iterations	fwd adjoint fwd + adjoint	linear nonlin-ear	Robertson & Langner, AE 1998, 4219 Elbern et al., JGR 1997, 15.967
Lagrangian	FWD	SRM	# sources	forward	linear	Maryon & Best, AE 1995, 1853
Lagrangian	BWD	SRM	# receptors	forward	linear	(prelim. work by Uliasz and Flesch et al.)

Eulerian models are run either in forward mode, or their adjoints in backward mode. These models are best suited for grid-to-grid calculations. Lagrangian particle models have the big advantage that they can simulate point releases much more accurately as they are not depending on a grid of a given scale. They would therefore be preferred for forward modelling of point sources or backward modelling of point measurements (this is the case in the CTBT context). However, backward-running Lagrangian particle models are just about to be established as a tool for SRM determination (Flesch and Wilson, 1995; Seibert, 2000)

Table 2 compares Eulerian and Lagranian methods. The conclusion is: If possible, use a Lagrangian model. Use the backward mode if you want to derive a gridded source. Use forward mode if you want to derive a point source whose location you know, or if there are only few possibilities for it.

4. Other related quantities

4.1 Fields of regard

As to my knowledge, the so-called “fields of regard” (FOR) as calculated by the IDC atmospheric transport software package represent source-receptor matrices. Each field is related to one receptor element (one station, one measurement interval) and contains the potential influence of source elements on a global grid for a certain release time (interval?). The wording used in the IDC manual, referring to ‘probabilities’, is misleading. The FOR has nothing to do with the probability of finding a source at a certain locations. Rather, the value of the FOR seems to indicates the magnitude of the concentration for a given source, or inversely, the source magnitude required to produce a specified concentration at the receptor.

One should also consider that information is not only contained in the FOR pertaining to a single observation where a suspect radionuclide has been detected. Zero measurements contain information on where a source should *not* be located, thus additionally constraining the potential source area. Making full use of the information available is hardly possible without a formal optimisation / inversion.

Table 2: *Advantages and disadvantages of the Eulerian and Lagrangian (forward / backward) approach in inverse modelling.*

Modell		**Advantage**	**Disadvantage**
Eulerian		moderate computational demands for typical resolutions straightforward to include nonlinear chemistry	limited vertical and horizontal resolution -> numerical diffusion and other problems point sources not well represented point receptor values must be interpolated from grid values
Lagran-gian	**general**	good numerical accuracy (infinite resolution during transport, finite resolution only when output is generated)	computational demand may be high, but depends strongly on model implementation and parameter choice
Lagran-gian	**FWD**	point source captured well good for many measurements and few source elements	point measurements interpolated from grid values, or computationally demanding kernel required
Lagran-gian	**BWD**	point measurements represented well good for many source elements and few measurements	point sources cannot be captured well except if their position is already known (kernel)

4.2 Adjoint tracer fields

Pudykiewicz (1998) has suggested to use so-called adjoint tracers for the source determination problem in the CTBT or similar contexts. He performed one backward calculation with an adjoint Eulerian model, releasing a pseudotracer at observation sites proportional to the measurements. It is claimed that the source location should be expected at the location of the maximum of the adjoint tracer field at the release time. For reasons similar to those discussed above, this is, however, not necessarily the case. Also, by not discriminating tracers from different sources, and using zero releases for measurement sites with zero measurements, information that could be utilised is lost.

This criticism does, of course, not mean that the adjoint approach in the context of a Eulerian model is useless. However, as pointed out in Section 3 it can only be used either as a means to calculate the SRM, or in an iterative approach to find the source term that minimises a given cost function.

5. Conclusions

The determination of SRMs is rather straightforward, and practically every dispersion model can be used for it. However, consideration should be given to the choice of the model and its implementation to optimise performance and accuracy according to the circumstances. The CTBT verification context, with a limited number of receptors and a very large number of potential source locations would favour backward calculations. However, if a potential source location has already been detected by a waveform technology, forward simulations would be more efficient and more accurate. Lagrangian particle models would probably be more accurate than Eulerian models in both cases.

The inversion of the SRM is less straightforward, especially as it may require additional assumptions. IDC software should thus contain tools to visualise combinations of subvectors of SRMs, and flexible tools for inversion. The usage of the inversion tools should, however, become a part of operational methods only after sufficient testing and clarification on how its results can be interpreted properly. It should also be possible to link SRM elements with respective measurements and thus obtain the source strength required to produce the observed concentration $x_i = m_{ij}^{-1} y_j$.

Inverse modelling of atmospheric trace constituents is an active area of research. The CTBT verification can give important stimuli to this research, whereas those in charge of the methods applied at the CTBTO/(P)TS should strive to utilise scientific advances.

Acknowledgements

This work is supported by the Austrian Federal Ministry of Education, Science and Culture (BMBWK).

References

Flesch, T. K. and J. D. Wilson (1995), Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol.* **34**, 1320–1332.

Høst, G. (1996), A statistical method for estimation of European sulfur emissions using EMEP monitoring data. EMEP / MSC-W, Note 3/96, 28 pp.

Menke, W. (1984), *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Orlando, 260 pp.

Nodop, K., W. Klug, A. Kulmala, H. V. Dop, J. Pretel, R. Addis, G. Fraser, G. G. F. Girardi, Y. Inoue, and N. Kelly (1998), ETEX: a European tracer experiment; observations, dispersion modelling and emergency response. *Atmos. Environ.* **32**(24), 4089 – 4094.

Pudykiewicz, J. A. (1998), Application of adjoint tracer transport equations for evaluating source parameters. *Atmos. Environ.* **32**(17), 3039–3050.

Robertson, L. and J. Langner (1998), Source function estimate by means of a variational data assimilation applied to the ETEX-I tracer experiment. *Atmos. Environ.* **32**, 4219–4225.

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 (eds.), *Inverse Methods in Global Biogeochemical Cycles*, pp. 147–154, AGU Geophysical Monograph Vol. 114, ISBN 0-87590-097-6, Washington.

Seibert, P. (2000), Inverse modelling with a Lagrangian particle dispersion model: application to point releases over limited time intervals. In: F. Schiermeier and S.-E. Gryning (eds.), *Proceedings of the Millenium International Technical Meeting on Air Pollution Modeling and its Application, May 2000*, pp. 284–291, American Meteorological Society.

Seibert, P. and A. Stohl (2000), Inverse modelling of the etex-1 release with a langrangian particle model. In: G. Barone, P. Builtjes, and G. Giunta (eds.), *Proceedings 3rd GLOREAM Workshop, Sept. 1999, Ischia , Italy*, pp. 95–105.

Stohl, A. (1998), Computation, accuracy and applications of trajectories – a review and bibliography. *Atmos. Environ.* **32**, 947–966.

Stohl, A., M. Hittenberger, and G. Wotawa (1998), Validation of the Lagrangian particle dispersion model flexpart against large-scale tracer experiment data, atmospheric environment. *Atmos. Environ.* **32**(24), 4245–4264.