Uncertainties in atmospheric dispersion modelling and source determination

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

In the context of the verification of the Comprehensive Test Ban Treaty based on the radionuclide monitoring network and atmospheric dispersion modelling, the uncertainties inherent to the model results and their appreciation in the interpretation process can be a highly critical matter. However, the issue of uncertainties in atmospheric dispersion modelling has so far been addressed by the scientific community in rather general terms only, and we are still far from established practices. This contributions presents a list of sources of uncertainty in dispersion modelling and the subsequent source determination together with preliminary ideas how they might be quantified in an operational context.[1]

2. Sources of errors in dispersion modelling

Models of atmospheric transport and dispersion are based on meteorological input, such as wind and temperature fields, additional parameterisations of quantities that are not directly available from routine numerical weather prediction (NWP) model output, such as turbulence parameters, and a numerical simulation algorithm. Accordingly, we can classify the error sources into the analysis (or forecast) errors, the parameterisation errors, and the numerical errors.

2.1 Analysis errors

In the CTBT verification context, *a posteriori* calculations done with analysed fields rather than forecast ones are of prime importance, so that we don't need to be concerned here too much about forecast errors (though very-short-range forecasts may be utilised to improve the temporal resolution, e.g. 3 hourly forecasts if analyses are available every 6 h). Analyses obtained from NWP models contain errors due to the limited coverage of the globe with observations (for example, analyses are in general less accurate on the southern hemisphere), due to inaccurate observations, and due to imperfections of the analyses schemes. Theoretically, it would also be possible that countries try to forge their radiosonde data transmitted on GTS if they want to conduct a clandestine test.

In our context, errors in the wind field which directly affect the transport are the most important ones, as they directly affect the transport patterns. Moderate errors in the wind direction can cause the indication of a completely wrong potential source area, especially if occuring in a flow pattern with strong deformation.

Errors in temperature fields can affect the calculation of diffusion parameters.

Errors in precipitation fields cause errors in the simulation of the wet scavenging of soluble gases and aerosol particles. Such errors may be severe. The problem is worsened by the fact that presently NWP centres do not analyse precipitation, so that one has to rely on short-term forecasts. These forecasts may suffer from spin-up problems. Another complication is that there can be a lot of sub-gridscale variability in precipitation, especially in convective precipitation.

2.2 Parameterisation errors

Parameterisation errors affect the following processes, with decreasing estimated significance: vertical turbulent mixing, wet scavenging, horizontal turbulent mixing, dry deposition.

[1] Corresponging to the informal nature of the workshop and the limited circulation of these proceedings, references in this paper represent only a small subjective subset of the relevant body of literature.

Let us have a closer look on the problem for the example of turbulent mixing (often also called diffusion). The diffusion in Eulerian dispersion models is calculated on the basis of the so-called K-theory, *i.e.* as diffusion with space- and time-dependent turbulent diffusion co-efficients K_z and K_h. In Lagrangian particle models, mixing by subscale atmospheric motions is simulated by adding a random component to the resolved-scale wind vector. These random components are constrained by certain statistical properties, typically their standard deviations σ_u, σ_v, σ_w, and respective Lagrangian time scales $T_{Lx,y,z}$. These turbulence parameters have to be derived from the wind and temperature fields (possibly using also turbulent kinetic energy, if available), using analytical relationships. These relationships (parametersations) have been derived from observations in field campaigns, from physical simulations (water tanks etc.), or from idealised theoretical and studies). They rely on the idealised, simple conditions. Thus, their application to arbitrary real situations will involve inaccurracies. A good parameterisation means that these deviations are in general (but not always!!) of acceptable magnitude.

Mixing by convective clouds is a very important process in certain climatic regions and at certain times, but many present transport and dispersion models don't have any convection parameterisation!

2.3 Numerical errors

The practical solution of transport and diffusion equations requires discretisation which is another source of errors. In Eulerian models, the most important source of numerical errors is the advection scheme. Eulerian advections schemes suffer from numerical diffusion, phase errors, and wiggles (over- and undershootings) if – overall more accurate – higher order schemes are used. Another major source of errors is the treatment of plumes whose size is too small to be resolved by the grid. Either a more or less complicated sub-model has to be used, or severe artificial diffusion takes place. Problems may also occur in strongly deformational flows, as plumes become more and more elongated and finally become much narrower than one grid distance—a problem relevant especially for long simulation times (Pudykiewicz and Koziol, 1998).

In Lagrangian particle models, particle trajectories can be calculated with high accuracy and there are no problems to represent small plumes—the reason why they are often considered superior. Numerical errors manifest in these models as statistical fluctuations of particle density, translating into fluctuations of concentration fields or time series, if the number of particles is too small, or becomes too small in a part of the plume.

Lagrangian puff models calculate advection in a Lagrangian framework, and represent the growth of puffs due to turbulent diffusion analytically. They suffer from specific problems, related to inhomogeneities (especially of the wind vector) within the puffs.

Finally, both types of models suffer from interpolation errors. Interpolation is necessary because NWP model output is typically available on time intervals of 1 to 6 hours whereas typical time steps are minutes up to an hour at most. This interpolation can be avoided in on-line dispersion codes, where the tracer transport is simulated within the NWP model. Lagrangian particle models (and Eulerian models, if they use a finer or otherwise different grid than the original NWP model) have to interpolate also in space to the particle positions. All these interpolations cause non-negligible errors. One should be aware, however, that coarse-grid Eulerian models, even if they do not experience interpolation errors, will suffer in their accuracy from lack of resolution.

3. Quantification of dispersion modelling errors

3.1 Analysis errors

There are different ways to estimate analysis errors. The most reliable would be the comparison with observations that have not entered the analysis. Obviously, this is possible only for limited research campaigns. There is also the problem that the easily accessible quantity, the deviation observation–analysis, is made up of three contributions which are difficult to separate: in addition to the analysis error, there are also measurement errors and errors due to lack of representativity of the measurement for the scale represented in the specific analysis.

Another possibility is to compare analyses made at different NWP centres (Cabrit et al., 1999). The problems with this approch are that the methods are not independent (e.g., centres use mostly the same observations), and that the inter-model deviations are not necessarily representative for the deviation model–reality (all models may make the same bad mistake, or most models may perform badly while one model is close to reality, etc.).

Finally, advanced data assimilation schemes themselves produce error estimates themselves. Probably this would be the most useful operationally available estimate, though here again we have the problem to separate different contributions to errors (model and measurement error).

Next, we have the problem to translate analysis errors in transport (source-receptor matrix) errors. As the analysis errors are correlated in space and time by complex structures, a trivial Monte Carlo approach makes not much sense. Ensembles provided by probabilistic NWP systems are also not suitable, because they contain a very specific selection of initial states, typically characterised by different fast-growing synoptic features. However, transport patterns between certain receptors and sources are a completely different matter. Also, new ensembles are provided only every 24 h, and there is no identity between members of two subsequent ensembles, so that they cannot used as a basis for longer-term transport simulations.

3.2 Parameterisation errors

Parameterisation errors are extremely difficult to diagnose. It is possible either in limited field campaigns, or by comparing different models. But the latter method must make sure that the same input fields, numerical methods and interpolation schemes are used, what is not generally the case, and it suffers from the basic problem of "who is right" from pure model *inter*comparisons as outlined above.

3.2 Numerical errors

There is a vast body of literature on the numerical errors of advection schemes, mostly based on analytical or idealised numerical tests. I am not aware of studies quantifying these errors in practical transport applications.

The uncertainty due to particle density in Lagrangian models can be calculated with statistical methods, but this is not normally implemented in these models.

Stohl et al. (1995) have investigated interpolations errors. They showed that horizontal interpolation from 1° (2°) to 0.5° causes relative errors of about 5% (15%) in the horizontal and 25% (65%) in the vertical wind. Similar results were obtained for temporal interpolation. They also showed that transport deviations in trajectories calculated with differently resolved grids are on the order of several 100 km after a few days. These findings illustrate that there is an urgent need for the IDC to obtain well-resolved input data for its atmospheric transport calculations.

3.3 Overall error quantification

Model intercomparisons, for example those conducted in the framework of CTBTO–WMO co-operation, give an estimate of the overall errors, as different RSMCs use different analyses, parameterisation and numerical schemes. However, this estimate can be biased in both directions: in the end, all models rely on the same observations and scientific knowledge and thus may share systematic errors; on the other hand, some models may be better than others, so that differences rather indicate errors of the worse than of better models (pure intercomparison problem, see also above).

Tracer experiments (e.g., ETEX, CAPTEX, ANATEX—see, e.g., Stohl et al. (1998)) or events such as the releases from Chernobyl, Algeciras, forest fires (Wotawa and Trainer, 2000), volcanoes and so on (also called tracers of opportunity) are very useful to study model performances, but they are limited to a few samples out of the infinite manifold of weather patterns (see the extreme difference in model performance between the first and the second ETEX release!(Nodop et al., 1998)).

4. Source determination errors

I am basing this discussion on the method of calculation and subsequent regularised inversion of a source-receptor matrix (SRM) (Seibert, 2000). Errors can enter into the inversion through the SRM, the observed data, and the regularisation. Errors related to the solution of the linear system of equations

(LSE) can be controlled effectively. The main problem is the fact that often the structure of the SRM matrix is ill-conditioned, so that small errors in the data or in the SRM itself can cause huge errors in the results. This is why we need the regularisation. However, the regularisation means a trade-off between a-priori knowledge or assumptions and information contained in the data. Bayesian methods allow to find the optimum trade-off and to indicate the uncertainty of the result, but only if the uncertainties of all the input values are really known. Even without formally reverting to Bayesian theory, error bars in each observation can be translated into error bars of the solution of the regularised LSE (Menke, 1984).

5. Conclusions

A variety of errors affect the results of transport calculations as well as source determination. The biggest challenge in their quantification is the quantification of the transport errors. More research, but also more operationally available uncertainty quantification is desirable and is likely to become a major topic in atmospheric sciences in the future.

In certain sensitive synoptic patterns, transport errors may lead to completely wrong calculated transport patterns. Careful analysis and interpretation by atmospheric scientists is therefore a necessity before drawing conclusions in the case of a disputed event.

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

References

Cabrit, B., J.-P. Issartel, J.-M. Guerin, F. Hourdin, and I. Balkanski (1999), Test parameters determination: inversion errors and wind field uncertainties. Proc. of the Informal Workshop on Radionuclides, Bruyeres-le-Chatel, Sept 28 to Oct 2, 1998. CEA, BP12, F-91680 Bruyeres-le-Chatel, France.

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. and A. Koziol (1998), An application of the theory of kinematics of mixing to the study of tropospheric dispersion. *Atmos. Environ.* **32**(24), 4227–4244.

Seibert, P. (2000), Methods for source determination in the context of the CTBT radionuclide monitoring system. In: *this volume*.

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.

Stohl, A., G. Wotawa, P. Seibert, and H. Kromp-Kolb (1995), Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. *J. Appl. Meteorol.* **34**, 2149–2165.

Wotawa, G. and M. Trainer (2000), The influence of Canadian forest fires on pollutant concentrations in the United States. *Science* **288**, 324–328.