3-D chemistry-transport model Polair: numerical issues, validation and automatic-differentiation strategy

Vivien Mallet, B. Sportisse

To cite this version:

Vivien Mallet, B. Sportisse. 3-D chemistry-transport model Polair: numerical issues, validation and automatic-differentiation strategy. Atmospheric Chemistry and Physics Discussions, European Geosciences Union, 2004, 4 (2), pp.1371-1392. hal-00327858

HAL Id: hal-00327858

https://hal.archives-ouvertes.fr/hal-00327858

Submitted on 8 Mar 2004

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.
3-D chemistry-transport model Polair: numerical issues, validation and automatic-differentiation strategy

V. Mallet and B. Sportisse

CEREA, Joint Research Laboratory, École Nationale des Ponts et Chaussées, EDF R&D, France

Received: 3 October 2003 – Accepted: 3 March 2004 – Published: 8 March 2004

Correspondence to: V. Mallet (mallet@cerea.enpc.fr)
Abstract

We briefly present in this short paper some issues related to the development and the validation of the three-dimensional chemistry-transport model Polair.

Numerical studies have been performed in order to let Polair be an efficient and robust solver. This paper summarizes and comments choices that were made in this respect.

Simulations of relevant photochemical episodes were led to assess the validity of the model. The results can be considered as a validation, which allows next studies to focus on fine modeling issues.

A major feature of Polair is the availability of a tangent linear mode and an adjoint mode entirely generated by automatic differentiation. Tangent linear and adjoint modes grant the opportunity to perform detailed sensitivity analyses and data assimilation. This paper shows how inverse modeling is achieved with Polair.

1. Introduction

Several 3-D chemistry-transport models (hereafter CTM) are now available and have proven to be efficient in many applications from passive-transport simulations to data assimilation with highly non-linear models. One may cite Cmaq (Byun et al., 1998), Chimere (Vautard et al., 2001), Eurad (Hass et al., 1995), etc. However, building a code flexible enough to be able to deal with any of those applications still remains a challenge.

Polair is an Eulerian three-dimensional chemistry-transport model developed at ENPC (École Nationale des Ponts et Chaussées). It is designed to handle a wide range of applications. Thus it has been used for passive transport by Issartel et al. (2003), for impact at European scale, for photochemistry by Sartelet et al. (2002) and mercury chemistry by Roustan et al. (2003). Developments are led to include size resolved aerosols and other chemical mechanisms. In this paper, we focus on photo-
chemistry.

Beyond academic simulations coming along with code validation and basic modeling work, validations were undertaken through comparisons with other models and with measurements provided by dedicated campaigns. Therefore fine analyses identified strengths and weaknesses of the model. Some analyses may be found in papers previously cited. This paper reports results obtained at regional scale and European scale. At regional scale, the intensive observation period (hereafter IOP) # 2 of the ESQUIF campaign (from summer 1998 to winter 2001, over Île de France) enables comparisons with measurements. At European scale, a simulation over four months (from May to August 2001) was compared to measurements from several countries.

A major advantage of Polair is the tangent linear mode and the adjoint mode that are entirely generated by automatic differentiation. The availability of those modes enables to get sensitivities with respect to most parameters and to put into practice virtually any data-assimilation method involving the derivative of some output variables with respect to some input parameters. In this paper, we explain the whole process.

The paper is organized as follows. We first describe some abilities of Polair and we point out some numerical issues. Then we report results from validations against measurements from the ESQUIF campaign (regional scale) and measurements collected for four months in 2001 over Europe. Finally we explain the way the tangent linear and adjoint codes are generated and used.

2. Model overview

As an Eulerian three-dimensional chemistry-transport model, Polair handles basic physical processes by integrating in time the following equation:

$$\frac{\partial c_i}{\partial t} + \text{div} (c_i \cdot V) = \text{div} (K \cdot \nabla c_i) + \chi_i(c) - D_i + E_i$$

where $i$ labels a chemical species, $c$ is a vector of chemical concentrations, $V$ is the wind vector, $K$ is the diffusion matrix, $\chi_i$ combines production and loss terms of chemi-
chemical reactions, $D_i$ is the deposition term (dry deposition and wet deposition – scavenging), $E_i$ stands for the emissions (surface and volumic emissions).

In our model, $V$ is prescribed (e.g. by offline meteorological forecasts). The diffusion matrix is assumed to be a diagonal matrix. Horizontal diffusion coefficients $K_{xx}$ and $K_{yy}$ are not well known and are assumed constant in time and space. The vertical diffusion coefficient is estimated according to a given parameterization. For instance, we usually use the parameterization proposed in Louis (1979).

Chemical reactions are modeled by $\chi_i$ which depends on species concentrations. For the time being, Polair has several chemical mechanisms:

- Mercury chemistry: simplified mechanism based on Petersen (1995);
- Aerosols: modal approximation for inorganic aerosols;
- Photochemistry (for ozone): several mechanisms are available among which RADM 2 (Stockwell et al., 1990), RACM (Stockwell et al., 1997), EURORADM (Schell, 2000), MOCA (Aumont, 1994), CBM IV (Gery et al., 1989).

Dry gaseous deposition $D_i$ is computed as in Wesely (1989) or as in Baer et al. (1992), and required land use coverage data may be provided by the USGS data base (http://edcdaac.usgs.gov/glcc/glcc.html). Wet deposition is parameterized as in Sportisse et al. (2002).

Emissions $E_i$ are generated from coarse emissions of a few emitting classes and for a few aggregated species. Speciation and aggregation stages are followed to estimate $E_i$ as in Middleton et al. (1990).

Equation (1) is relevant at many scales, which lets a CTM be relevant from regional scale to continental scale. Since Polair can easily handle several chemical mechanisms, different applications may reasonably be led at those scales.

The modularity of Polair has been strengthened by computing all parameterizations in preprocessed steps. As such, Polair only is a numerical platform for solving
advection-diffusion-reaction partial differential equations. The physical parameterizations are computed by the library AtmoData (Mallet, 2004) and the Fortran files describing chemical kinetics are automatically generated by an appropriate preprocessor (SPACK: Simplified Preprocessor for Atmospheric Chemical Kinetics, Djouad et al., 2002).

3. Some numerical issues

Numerical issues have received a special attention because of computational requirements (specially for data assimilation). We focus on the integration of chemistry because of its prominent computational cost.

3.1. Time integration

Concentrations of all chemical species over all cells are coupled through Eq. (1), which can lead to very high computational requirements. To avoid prohibitive computations, the equation is split mainly into three parts: advection, diffusion and chemistry. Processes are integrated in a sequential way. For instance, a first-order splitting method would integrate the advection over the time step $\Delta t$, the diffusion over the same time interval and then the chemistry over the same interval. It has been shown (Sportisse et al., 2000) that this order (advection, diffusion and chemistry) is preferred. A Strang splitting is available in Polair as well.

The advection numerical scheme is the third order Direct Space Time scheme (advocated in Spee, 1998) which takes advantage of a Koren-Sweby flux limiter. The diffusion is integrated by the second-order Rosenbrock method described below.

3.2. Integration of chemistry

It is well known that equations to be solved (see Eq. 1) introduce a wide range of characteristic timescales since chemical reactions have strongly different reaction rates.
Let us consider the Jacobian matrix $\mathbf{A}$ defined by $\mathbf{A}_{ij} = \frac{\partial \chi_i}{\partial c_j}$. If the dependence of $\chi$ with respect to time is locally neglected, a linearized version of chemical kinetics is:

$$\frac{\partial \mathbf{c}}{\partial t} \sim \chi(c(t = 0)) + \mathbf{A} \mathbf{c}$$

The characteristic timescales of the dynamical part are therefore determined by the eigenvalues of $\mathbf{A}$. Figure 1 shows a typical eigenvalues distribution. Such a distribution justifies the need for a numerical solver dedicated to stiff equations.

In practice, we use the second order Rosenbrock method (see Verwer et al., 1999):

$$c_{n+1} = c_n + \Delta t b_1 k_1 + \Delta t b_2 k_2$$

(2)

where

$$\begin{cases} k_1 = \chi(t_n, c_n) + \Delta t \gamma \mathbf{A} k_1 \\ k_2 = \chi(t_n + \alpha_{12} \Delta t, c_n + \alpha_{21} \Delta t k_1) + \Delta t \gamma_{21} \mathbf{A} k_1 + \Delta t \gamma \mathbf{A} k_2 \end{cases}$$

(3)

and $\alpha_{12}$, $\alpha_{21}$, $\gamma_{21}$ and $\gamma$ are prescribed coefficients and $\Delta t$ is the time step.

However this scheme is well suited in case of autonomous equations. Because of photochemistry and its dependence upon time (solar radiations), the equation to be solved is non-autonomous. As a consequence, the positivity of the solution may not be preserved and negative concentrations are set to zero in order to avoid instabilities. A solution to retain positivity could also be to reduce the time step. Another solution is the stabilized Rosenbrock method (see Djouad et al., 2002) where Eq. (3) is replaced with:

$$\begin{cases} (\mathbf{I} - \gamma \Delta t \mathbf{A}) \cdot k_1 = \chi(t_n, c_n) + \gamma \Delta t \partial_t \chi(t_n, c_n) \\ (\mathbf{I} - \gamma \Delta t \mathbf{A}) \cdot k_2 = \chi(t_{n+1}, c_n + \Delta t k_1) - 2k_1 - \gamma \Delta t \partial_t \chi(t_n, c_n) \end{cases}$$

(4)
which takes into account the time dependence of the production term $\chi$. From our tests, the production-term non-autonomy and the time step (with which simulations are performed) are so that the non-autonomous scheme is not useful. However clipped masses are very small as compared to the whole pollutant masses. According to our tests, they can be neglected.

3.3. Cost reduction of the Rosenbrock integration

The main cost of the Rosenbrock method comes from the computations of $k_1$ and $k_2$ in Eq. (3) which imply to solve a similar linear system twice. Its size is $N \times N$ where $N$ is the number of chemical species. We use an LU factorization of $J = I - \gamma \Delta t \Lambda$. An important feature of $J$ is its sparsity: about 85% of matrix elements are zero. This is the reason why the method proposed by Sandu et al. (1996) can be very efficient. We permutate species so that the LU factorization of $J$ remains sparse. This can be achieved by many algorithms, one of which is the Markowitz algorithm. Moreover indices of non-zero entries are known a priori. Non-zero entries of $J$ are deduced from species involved in chemical reactions, and non-zero entries of its LU factorization can be determined along with the factorization itself. Since indices of non-zero entries are known, it is possible to build tailored factorization and solver algorithms. They are usual algorithms from which useless computations (involving zero entries) were removed. Finally, we produce a code without indirect addressing (as in the Harwell-Boeing storage) also because indices of non-zero entries are known a priori.

This method dramatically reduces computational costs. For instance, the integration of the chemical mechanism RADM 2 (with 61 species) was speeded up by a factor of 3.6.
4. Validations

Two validations were conducted in order to assess model capabilities in non-academic cases. The first validation takes place over Île de France, i.e. at regional scale. The second one tests the model at continental scale and over a long period (four months).

4.1. Validation at regional scale

The ESQUIF campaign investigated photo-oxidant pollution and led twelve intensive observation periods. IOP #2 occurred in August 1998 with low wind speeds (stagnation), a clear sky and high temperatures. This highly polluted episode has to be correctly simulated by a model that aims at useful forecasts. The IOP #2 was simulated over Île de France from 7 to 9 August 1998.

The 150 km × 150 km domain is divided into 25 cells along x (6 km), 25 cells along y (6 km) and 9 cells along z (cell-centers altitudes being 15 m, 90 m, ... up to 2525 m). The top altitude of the last cell is 2800 m which seems reasonable according to the height of the boundary layer. We have taken a fixed time step of 600 s.

The main chemical mechanism was RADM 2 (see Stockwell et al., 1990), with 61 species and 157 reactions. The same simulation was performed with the chemical mechanism MOCA, with 66 chemical species and 126 reactions. However only results from RADM 2 will be analyzed here since both chemical mechanisms gave very similar results. Initial and boundary conditions come from output concentrations of the 3-D CTM Chimere (Vautard et al., 2001). Dry deposition velocities are assumed to be constant.

A simulation takes less than 2 min per day on an Intel Pentium IV, 2 Ghz, and with the Intel Fortran Compiler.

Results for ozone at four stations are shown on Fig. 2. They were linearly interpolated from output concentrations of Polair. Figure 2a – station Paris 18 – shows a reasonable agreement with measurements inside Paris. Peaks are well forecast and grow from one day to the other due to stagnant conditions. Figure 2b – station Ram...
bouillet – demonstrates a good behavior for the two last days. However the first day is not correctly simulated. Rambouillet was located in Paris plume (south west) and was highly dependent on simulated concentrations inside Paris. However initial conditions at Rambouillet and along the plume come from a continental model, which means that the first day is constrained by erroneous concentrations that cannot resolve the plume. Figure 2c – Aubervilliers – shows overestimated concentrations at a station close to Paris 18 but outside Paris (north east) and its plume. The reason for this overestimation is not identified. This may come from lack of biogenic emissions or from boundary conditions. Biogenic emissions were not included because there were simply not available in Polair (at the time – it is not the case anymore). Moreover, a chemical mechanism like RADM 2 is not perfectly suited to deal with biogenic emissions. One may use RACM instead. Figure 2d – Montgé en Goële – is a rural station which is not in the plume. The first day is not well simulated because of initial conditions. The two other days show again an overestimation, as for station Aubervilliers.

Figure 2 summarizes what is observed at other stations. The first day suffers from initial conditions. As for the other days, results at stations within Paris and its plume are reasonably close to measurements. Those results are driven by rather fine data like anthropogenic emissions. Outside Paris and its plume, results are not as good as expected and are overestimated. At stations above the first level (30 m), improvements are needed as well.

To assess the whole results, the root mean square between computed concentrations and measurements is considered for all stations and over the three days. The mean over all stations of the root mean square is 39 $\mu$g·m$^{-3}$, for an episode with high ozone concentrations. The correlation is 86%, which indicates a good time evolution for Polair outputs.

Improvements should be obtained with better input fields. Meteorological data could be processed in a better way, deposition velocities should be computed thanks to Wesely (1989) or Baer et al. (1992) which are now available in the code, biogenic emissions should improve results specially outside the plume.
4.2. Validation at continental scale

At continental scale, a simulation over more than four months (from May to August 2001) demonstrates the validity of the model. This period gathers interesting episodes and is thus a complete validation in summer conditions.

The domain is \( (40.25^\circ \text{N}, 10.25^\circ \text{W}) \times (57.25^\circ \text{N}, 22.75^\circ \text{E}) \), with \( 0.5^\circ \times 0.5^\circ \) cells, namely 33 cells along latitude and 65 cells along longitude. There are only 5 cells along \( z \) whose centers are 25 m, 325 m, 900 m, 1600 m and 2500 m. The top height of the last cell is 3000 m. The time step is 600 s, even if a larger time step should be acceptable. The chemical mechanism is RACM.

The purpose of this section is not to get into details about this simulation. However let us review quickly input data. Meteorological data comes from the ECMWF, anthropogenic emissions are provided by the EMEP database, biogenic emissions are estimated as in Simpson et al. (1999), initial and boundary conditions are extracted from Mozart-2 (Horowitz et al., 2003) output data. Vertical diffusion is estimated by the Louis parameterization (Louis, 1979). Deposition velocities are computed according to Wesely (1989).

Measurements from 242 stations (assumed to be background stations) are available. The mean of root mean squares at those stations is 29.7 \( \mu \text{g} \cdot \text{m}^{-3} \) (computed with all measurements, i.e. hourly measurements) and is 24.7 \( \mu \text{g} \cdot \text{m}^{-3} \) for concentrations at 15:00 UTC. The correlation is 60\% (all measurements) and is 66\% for concentrations at 15:00 UTC. Those preliminary results show a good agreement with measurements.

Figures 3 and 4 show ozone peaks over the four months at two stations where root mean squares (25.2 \( \mu \text{g} \cdot \text{m}^{-3} \) and 24.5 \( \mu \text{g} \cdot \text{m}^{-3} \) respectively) are close to the mean of all root mean squares (24.7 \( \mu \text{g} \cdot \text{m}^{-3} \)). Figure 5 shows a scatter plot of measurements of ozone concentrations at 15:00 UTC versus concentrations computed by Polair.

Those preliminary results seem good enough to let one be confident in the model. Compared to other models in very similar conditions (e.g. through the Pioneer forecasts – http://euler.lmd.polytechnique.fr/pioneer/), the results are satisfactory.
They will enable one way nesting and therefore should improve simulations at regional scale (by providing boundary conditions from the same model and with the same chemical mechanism). Moreover they allow to perform meaningful data assimilation.

5. Inverse-modeling strategy

Polair was designed to be well suited for data assimilation, notably thanks to the availability of a tangent linear mode and an adjoint mode. In this section, we provide details about this feature and we present an application.

5.1. Tangent linear and adjoint modes

Let us consider Polair as a function that takes input parameters (meteorological fields, emissions, etc.) and returns concentration fields. If one can get derivatives of Polair output fields with respect to input parameters, one can get sensitivities of output concentrations to input parameters and one can build inverse modeling experiments or data assimilation experiments involving variational methods. Derivatives of Polair would therefore be of high interest.

Now, Polair can be automatically differentiated, i.e. one can get a code that returns derivatives of output concentrations (as computed by Polair) with respect to input parameters. Written in Fortran 77, it is differentiated by Odyssee (Faure et al., 1998), which provides the tangent linear code and the adjoint code of Polair. From those codes, derivatives (previously described) are available.

No modification in Polair is required. However, to enforce computational efficiency, we only replace the differentiated code of the LU factorization and the LU solver. To get such a simple procedure, the code was carefully built. For example, input/output operations (which are not differentiable) and numerical integration have been separated.

Moreover the use of the differentiated code is straightforward since input/output subroutines remain relevant. Loosely speaking, the numerical integration over one time
step is replaced by its differentiated code and the whole structure around it doesn’t change. In practice, it means that we can easily get the differentiated code with respect to virtually any parameter and for any simulation. Notice that absolutely all processes are differentiated: chemistry, diffusion, advection, deposition, etc.

Actually the main step is to validate the differentiated code by comparison with finite differences. This is performed with simulations involving all processes but at low costs (e.g. a few cells are enough). A derivative $D(\varepsilon)$ is computed by finite differences (i.e. by computing the difference between a reference simulation and a slightly perturbed simulation) and compared to the differentiated-code output $d$. As the perturbation $\varepsilon$ becomes smaller, finite differences should return a result closer to the differentiated-code result $d$. If the perturbation is too small, round-off errors destroy finite differences accuracy. A relevant test is therefore to compute the ratio between the finite-difference result and the differentiated-code result as function of the perturbation. Figure 6 displays results of such a (successful) test.

The computational cost of the differentiated code is more or less a function of the forward code cost. The tangent linear code costs twice as much as the forward code, and the adjoint code cost is between 4.5 and 7 times the forward code cost.

5.2. Application

An application of such a code may be sensitivity analysis. One may consider an output variable (e.g. a 3-D field of ozone concentrations, or a given cost function) and may need the derivative of this variable with respect to a set of parameters. This could be useful in order to choose relevant parameters for data assimilation: the more a variable is sensitive to a given parameter, the more relevant the parameter is. Such sensitivity analyses only provide local sensitivities. Nevertheless if several studies are performed in different situations, a very detailed analysis can be led and can provide results that encompass enough cases to describe global sensitivities.

We performed, at both regional and continental scales, sensitivity analysis of ozone, nitrogen oxide and nitrogen dioxide concentrations to emissions of many chemical
species. Analyses of results are beyond the scope of this paper.

As soon as relevant parameters are chosen (with or without the use of the differentiated code), inverse modeling and data assimilation (4-D variational method) can be led thanks to the adjoint code. Currently, inverse modeling of emissions by use of the adjoint code is under progress.

6. Conclusions

We have summarized the main features of the chemistry-transport model Polair. Two simulations at regional and continental scales have shown a reasonable agreement with measurements. Thanks to those validations and the availability of differentiated versions of Polair, sensitivity studies and data assimilation can be performed. The main future works will be devoted to aerosol modeling and air-quality ensemble forecast.

Acknowledgements. We thank I. Coll (LISA, Université Paris 12) for her help (simulation at regional scale), R. Vautard (LMD, École Polytechnique) and C. Honoré (Ineris) for their help (simulation at European scale), Airparif (monitoring air quality in the Paris region) for its emission data base over Île de France, and all monitoring networks that have made comparisons with measurements possible at European scale.

Edited by: J. Brandt

References


Byun, D. W., Young, J., Gipson, G., Godowitch, J., Binkowski, F., Roselle, S., Benjey, B., Pleim, J., Ching, J. K. S., Novak, J., Coats, C., Odman, T., Hanna, A., Alapaty, K., Mathur, R.,


1374, 1379, 1380
Fig. 1. Absolute values of eigenvalues of a Jacobian matrix coming from chemical mechanism RADM 2. The y-axis is a log-scale. Eigenvalues are strongly spread, from $10^9$ to $10^{-11}$ and smallest eigenvalues are omitted because of numerical limitations.
Fig. 2. Ozone simulated (from 7 to 9 August 1998) concentrations at: (a) Paris 18, (b) Rambouillet, (c) Aubervilliers and (d) Montgé en Goële.
Fig. 3. Ozone peaks (15:00 UTC) for each day of the simulation (from 27 April to 31 August 2001) at a station in France (Neuilly) where the root mean square is 25.2 µg m$^{-3}$. 
Fig. 4. Ozone peaks (15:00 UTC) for each day of the simulation (from 27 April to 31 August 2001) at a station in the Netherlands where the root mean square is 24.5 µg m⁻³.
Fig. 5. Scatter plot of measurements (at all stations) of ozone concentrations at 15:00 UTC versus concentrations computed by Polair.
Fig. 6. Comparison between the derivative $D(\varepsilon)$ as estimated by finite differences and the derivative $d$ computed by the automatically differentiated code. The ratio tends to 1 until round-off errors destroy finite-difference accuracy.