Inverse Modeling of NO_x Emissions at Regional Scale over Northern France. Preliminary Investigation of the Second-Order Sensitivity

Denis Quélo, Vivien Mallet and Bruno Sportisse

CEREA, Research Center for Atmospheric Environment, Joint Laboratory École Nationale des Ponts et Chaussées and Électricité de France, France

Abstract. The purpose of this article is to perform the inverse modeling of emissions at regional scale for photochemical applications. The case study is the region of Lille in the Northern France for simulations in May 1998. The underlying Chemistry-Transport-Model, Polair3D, has been validated with one year of model-to-observation comparisons over Lille. Polair3D has an adjoint mode, which enables inverse modeling with a variational approach. A sensitivity analysis has been performed so as to select the emission parameters to be modified in order to improve ozone forecasts. It has been shown that inverse modeling of the time distribution of nitrogen oxides emissions leads to satisfactory improvements even after the learning period. A key issue is the robustness of the inverted emissions with respect to uncertain parameters. A brute-force second-order sensitivity analysis of the optimized emissions has been performed with respect to other parameters and has proven that the optimized time distribution of NO_x emissions is robust.

1. Introduction

Emission inventories used in air pollution modeling admit a large range of uncertainties [*Hanna et al.*, 2001, for instance]:

1. the spatial distribution of emissions is not always well known and may be highly heterogeneous;

2. the time distribution of emissions is strongly related to variable parameters (such as traffic conditions, biogenic activity);

3. the chemical distribution is also uncertain: the relations between the chemical species given by the emission inventories, the "real chemical species" and the "model species" (the species described in Chemistry-Transport-Models) are often questionable.

A growing field is then logically the inverse modeling of emissions (more precisely: of parameters related to the emissions) on the basis of a combined use of model outputs and observational data (provided by monitoring networks). These topics belong to the larger domain of data assimilation.

Moreover there are further reasons to support these approaches. One may be interested in estimating the emissions of a given sector or of a given country in order to check the fulfillment of a regulatory agreement. This is typically the case for the gases implied in the Greenhouse effect at global scale or for the pollutants regulated by the Long Range Transport of Air Pollution Protocol (LRTAP) over Europe. Inverse modeling of emissions is also an interesting tool for such applications.

An increasing number of works has been devoted to these topics in recent years. At global scale, passive tracers or weakly reactive species, such as CO or CH_4, have already been studied. One can refer for instance to the work of *Kaminski* [1998]; *Bergamaschi et al.* [2000]; *Bousquet et al.* [1999]. In the case of linear tracers, such as radionuclides, many methods have already been proposed, following the Chernobyl accident and the ETEX campaign [*Hourdin and Issartel*, 2000].

The situation is quite different for reactive Chemistry-Transport-Models, as the dependence of concentrations on emissions is nonlinear. Moreover the models are characterized by high-dimensional systems (whose dimension is given by the number of chemical species in the chemical mechanism). One can refer to the work of *Elbern et al.* [2000] for academic studies and to *Elbern and Schmidt* [2001]; *Mendoza-Dominguez and Russel* [2001]; *van Loon et al.* [2000] for examples at continental scales. A few works have been devoted to the inverse modeling at regional scales: we can refer for instance to *Chang et al.* [1997] for the inverse modeling of biogenic isoprene emissions over Atlanta with the use of a Kalman filter. Another work is *Mendoza-Dominguez and Russel* [2001] with a linearized method applied to Atlanta, as well.

The purpose of this paper is to perform inverse modeling of emissions for air pollution applications at regional scale. A key issue, which is not often investigated, is the robustness of the inverted parameters: how to be sure that the results are not only "fits" of the model outputs to the data ? What is the quality of the new emission parameters ? How sensitive are these parameters with respect to other uncertain parameters (supposed to be known) ?

We have investigated these issues with an application to northern France, over the region of Lille. A comprehensive 3D Chemistry-Transport-Model, Polair3D [*Boutahar et al.*, 2004], has been validated by comparisons to measured data over one year [1998 in *Quélo*, 2004]. A sensitivity analysis has also been performed in order to choose the relevant parameters for inverse modeling. The choice was made to perform inverse modeling of the time distribution of NO_x emissions. The emissions of NO_x are the emissions that have, at first-order, the greatest impact on ozone in this case and their time distribution is not well known (contrary to the spatial distribution, an exhaustive emission inventory having been built in the framework of the French Research Program PREDIT, a program for research, experimentation and innovation in land transport).

0148-0227/05/$9.00

Inverse modeling has been performed through variational methods, Polair3D having an adjoint mode. The learning database is composed of data from the week of 11-15 May. The use of the modified emission inventory improved the simulated results of the two following weeks. Moreover the robustness of the inversion has been investigated.

This paper is organized as follows. The second section is devoted to the general presentation of the case study. In the third section, some preliminary tests are made in order to assess the sensitivity of model outputs with respect to emission data. The cost function, that measures the discrepancy between observations and model outputs, is also studied. In the fourth section, the inverse modeling approach is presented and twin experiments (on the basis of numerical data) are performed in order to validate the numerical models. The impact of uncertainties (for instance model errors) is investigated. In the fifth section, the real case is studied with the inversion of a time distribution for NO_x emissions. A posteriori verifications and robustness tests are also performed in order to assess the quality of the optimized parameters.

2. Case Study

Lille is part of a dense urban area in Northern France that includes many intermediate cities. Therefore the majority of pollution comes from local anthropogenic activities. Pollution may also result from nonlocal sources since Lille is sometimes located in the plume of highly polluted areas (Paris, Ruhr, London). In order to take into account this plume, a simulation at European scale is performed first and the simulation domain over Lille is nested in it.

2.1. Brief Overview of the Chemistry-Transport-Model: Polair3D

Polair3D [*Boutahar et al.*, 2004] is a comprehensive 3D Eulerian Chemistry-Transport-Model developed at CEREA (laboratory at École Nationale des Ponts et Chaussées and the Research and Development Division of Électricité de France). It is one part of the Polyphemus modeling system (*Mallet et al.* [2005], also developed at CEREA and available under the GNU General Public License at `http://www.enpc.fr/cerea/polyphemus/`), notably devoted to impact studies, forecasts and data assimilation for the atmospheric dispersion of chemical species and radionuclides. Within this system, Polair3D is mainly responsible for the time integration of the chemistry-transport equation. It includes several chemical mechanisms, including RACM [*Stockwell et al.*, 1997] which was chosen for this study. The other components of Polyphemus provide the input fields to Polair3D (meteorological fields, deposition velocities, etc.), computed using relevant physical parameterizations.

One of these components is the library AtmoData [*Mallet and Sportisse*, 2005] which gathers the physical parameterizations. Thanks to this library and the programs available with it (themselves part of Polyphemus), the following fields were computed for this study:

1. the meteorological fields extracted from ECMWF data;

2. the vertical diffusion coefficients computed with Louis' parameterization [*Louis*, 1979];

3. the cloud attenuation computed in the same way as advocated in *Chang et al.* [1987]; *Madronich* [1987];

4. the deposition velocities with Wesely's parameterization [*Wesely*, 1989];

5. the anthropogenic emissions from the EMEP inventory, following *Middleton et al.* [1990], at European scale; at regional scale, the emissions are generated as explained in section 2.3;

6. the biogenic emissions computed as advocated in *Simpson et al.* [1999];

7. the boundary conditions (for the European simulation) extracted from a Mozart 2 [*Horowitz et al.*, 2003] simulation over a typical year.

All these fields, that appear in the chemistry-transport equation, are then available to Polair3D which in turn integrates the equation in time with efficient numerical schemes. It uses a first-order splitting method in which the advection is integrated first, then the diffusion and finally the chemistry. The advection scheme is a third-order direct-space-time scheme with a Koren flux limiter and is advocated in *Verwer et al.* [1998]. The diffusion and the chemistry are both integrated with a second-order Rosenbrock method which is suitable for stiff problems and whose implicitness enables the use of large time steps (in this study, 600 s at both European scale and regional scale). As for chemistry, to enforce the computational efficiency, the sparsity of the Jacobian matrix involved in the Rosenbrock method is taken into account as advocated in *Sandu et al.* [1996].

The simulations performed with Polair3D have proven to be reliable. At European scale, the model has been validated notably over the year 2001. Details about this validation may be found in *Mallet and Sportisse* [2004]. For instance, the comparison with measurements from May to August 2001 of ozone peaks at 242 stations ($27,000$ measurements) gives a root mean square of $22.7 \mu g \cdot m^{-3}$ and a correlation of 72.7%. At regional scale, the results are also satisfactory as shown in the following sections.

One should note that the modeling system Polyphemus provides data sets and parameterizations consistent with the current knowledge in physics. Numerical adjustments that are not supported by physics were discarded even if they could lead to better results. This is the only means to ensure that the inverse modeling of emissions makes sense, i.e. that the retrieved emissions have a chance to be closer to the real emissions.

The last point to be emphasized is the availability of a tangent linear mode and an adjoint mode of Polair3D with respect to virtually any input parameter, including the emissions. This feature is permitted due to automatic differentiation [*Mallet and Sportisse*, 2004].

2.2. Domain

The domain covers the region of Greater Lille, over a 21 km × 24 km domain. The center of the domain is the city of Lille. It is discretized with a 1 km × 1 km horizontal grid and 9 vertical levels ranging from the ground to 3000 meters in order to include the atmospheric boundary layer. The height of the first layer is 30 m and the thickness of the other layers ranges from 120 m to 510 m.

2.3. Emissions over Lille

The anthropogenic emissions come from several databases:

1. EMEP inventory for VOCs, NO_x, SO_2 and CO, which is the results of a "top-down" approach and is available in annual totals over a 50 km × 50 km grid for each activity sector (SNAP – Selected Nomenclature for Air Pollution);

2. traffic emissions delivered by the CETE institute (Centre d'Étude Technique de l'Équipement) over the Lille regional administrative area on the basis of the road locations, the distribution of vehicle categories and the emission factors from the standard COPERT III European methodology;

3. total annual emissions of major industrial sources collected by the DRIRE institute (Direction Régionale de l'Industrie, de la Recherche et de l'Environnement): within the Lille area, 20 point sources are taken into account including production processes and waste treatment.

Emissions from the EMEP inventory are replaced with the detailed data provided by the two last databases whenever possible. A large part of the NO_x emissions come from the traffic (about 60%) and are therefore well described by our inventory. This is highly valuable according to the sensitivity of photochemical concentrations to the NO_x emissions.

The EMEP annual totals are first mapped to the simulation grid. Because of the coarse resolution of the EMEP inventory, the spatial distribution is mainly determined by the land use coverage in order to associate the major part of the emissions with urban areas. Thus, within an EMEP cell, the emissions are distributed so as to give an appropriate weight to the urban areas, the forests and the other areas (respective weights: 12, 1.6 and 1). The two other databases are already accurately distributed and their horizontal grid matches the simulation domain.

The emissions are then vertically distributed to take into account the stack heights and the elevation due to the high temperature at release time.

Annual EMEP emissions and point sources are distributed monthly, weekly and hourly according to coefficients provided by GENEMIS (Eurotrac-2 subproject – `http://www.gsf.de/eurotrac/`) for each emission sector. The time distribution of road traffic emissions is computed on the basis of daily and hourly coefficients derived from traffic activity measurements.

It is assumed that the NO_x emissions are composed of NO (10%) and NO_2 (90%). The NMVOC emissions are computed following *Middleton et al.* [1990].

2.4. Monitoring Network

The locations of the stations of the monitoring network AREMA are plotted in figure 1 for ozone and in figure 2 for NO_x (one measurement of NO and NO_2 per station). The network includes urban and suburban stations and delivers hourly measurements.

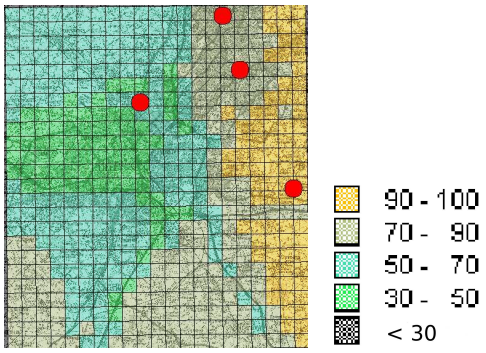

Figure 1. Monitoring network (red dot) and simulated daily concentrations for ozone over Lille, 11 May 1998 (μgm^{-3}).

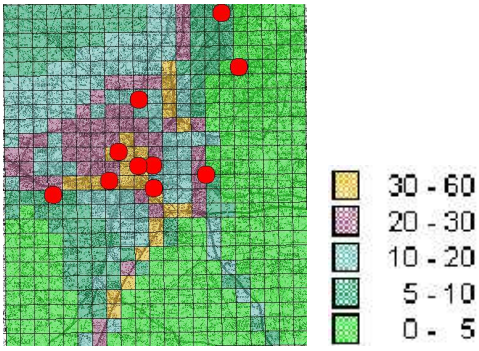

Figure 2. Monitoring network (red dot) and simulated daily concentrations for NO over Lille, 11 May 1998 (μgm^{-3}).

2.5. Validation over Lille

Polair3D has been used in order to simulate air quality over Lille for the year 1998. The boundary conditions have been provided by continental runs of Polair3D. We refer to *Quélo* [2004] for a more detailed description of the validation.

The model outputs have been compared to measured data provided by the local monitoring network, AREMA. Output concentrations have been interpolated to the locations of the monitoring stations. Statistical measures (root mean square, bias, correlation) have been computed with hourly data for three species: O_3, NO_2 and NO. Polair3D has shown a satisfactory agreement between simulated concentrations and observations [see *Quélo*, 2004]. In particular, the correlation for NO_2 is above 47% at all monitoring stations except one, which proves that Polair3D reproduces well the spatio-temporal variability of NO_2 concentrations.

2.6. Setup of the Inverse Modeling Case

The setup of the modeling case is the same as for the forward simulations. The full simulation system has been used, without any limitations in the physics or in the numerical schemes.

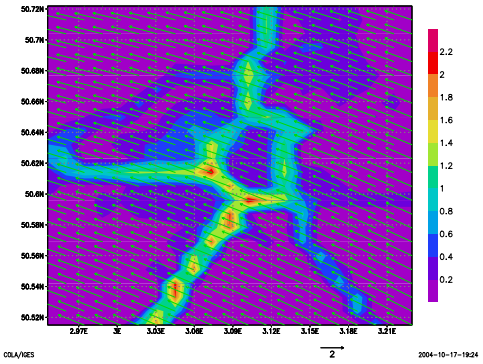

Figure 3. Average surface emissions for NO in $\mu g \cdot m^{-2} \cdot s^{-1}$ and average wind at ground in $m \cdot s^{-1}$. 11 May 1998.

We have chosen to study the three weeks after 11 May.

We have focussed on the month of May. The average wind field at the ground is plotted in figure 3 for one day (11 May). For this day, the average wind velocity is of the magnitude 2 $m \cdot s^{-1}$, which corresponds to a residence time of 3 hours for the pollutants in the simulation domain. This situation is quite representative for this month.

3. Some Preliminary Tests: First-Order Sensitivity Analysis

3.1. Sensitivity Analysis of the Cost Function

The purpose of this subsection is to investigate the sensitivity of a cost function, that describes the discrepancy between model outputs and observational data, with respect to input parameters for Polair3D. Similar studies have already been performed but they have not been devoted to cost functions (for instance *Segers* [2002] over England and *Menut* [2003] over Paris). This is a key step in order to assess the feasibility of inverse modeling.

3.1.1. A Few Notations

In the following, we consider that the model outputs of Polair3D are of the form:

$$c = f(k) \tag{1}$$

c is a vector of space- and time-distributed chemical concentrations and is called the *state vector*. k is the vector of all the input parameters to the Chemistry-Transport Model (see below for examples).

The observations may be compared to c thanks to an observation operator, which is usually written as H. In our case (ground observations), H is a projection matrix, which maps the vector c to the values of several chemical concentrations at the monitoring stations (at given spatial locations and given dates). Hc is therefore the vector of observations deduced from the state c: $(Hc)_i$ is the i-th observation, where i is an index labelling time, space and chemical species. obs_i represents the observations, i.e. the measured concentration for species O_3, NO and NO_2 (measured by the monitoring network AREMA).

The cost function is defined in order to estimate the discrepancy between the observations (obs) and the numerical results provided by the model ($Hf(k)$). It is usually written in the following form:

$$J = (Hf(k) - obs)^T R^{-1} (Hf(k) - obs) \tag{2}$$

where R is the so-called covariance matrix of the observation error.

In order to perform a sensitivity analysis, we multiply the input parameters k by a scalar numerical parameter α. Typically, α ranges from 0.5 to 1.5 in order to model perturbations of magnitude 50% for k. The reference case is given by $\alpha = 1$. We then write the cost function as a function of α with given values of the input parameters k (given by the reference case):

$$J(\alpha) = (Hf(\alpha k) - obs)^T R^{-1} (Hf(\alpha k) - obs) \tag{3}$$

Hereafter, we assume that the observations are uncorrelated and that the observation error variances are the same for all observations. In practice we can therefore take $R = I$, the identity matrix. The cost function is then a spatial and time average, over the monitoring network and over one day (for this section), respectively.

The evolution of J with respect to the perturbation parameter α has been computed for several input parameters k:

1. the lateral boundary conditions provided by the continental simulation for NO, NO_2, O_3 and the remaining species (figure 4);
2. the vertical eddy coefficient K_z (figure 5);
3. the kinetic rate of the reaction $O_3 + NO \rightarrow NO_2$ (in order to quantify the segregation effects on chemical kinetics, figure 6);
4. the emission fluxes for the primary pollutants NO, NO_2, CO, SO_2 and VOC (figure 7);
5. the dry deposition velocities for NO_2, O_3 and the remaining species (figure 8);
6. the attenuation coefficient for photolysis (parameterizing the effects of clouds; figure 9).

3.1.2. Results

The simulations have been performed over one day (11 May). The results depend on the chosen day but are representative of the typical sensitivity levels (results not reported here). In the following, we have plotted the sensitivity with a constant variation of α ranging from 0.5 to 1.5, which may not correspond to the actual uncertainties for the whole parameters. Even if the magnitude of these uncertainties may be more or less known (see *Hanna et al.* [2001] for instance), we have chosen to quantify the impact of similar perturbations in the inputs. Notice that the scale of each figure has been adapted to the values of the cost function.

As ozone is a regional/continental pollutant, it is logical to get a large impact of ozone boundary conditions. On the other hand, the impact of NO_x boundary conditions is much weaker (figure 4).

The dependence on K_z at the first level (30 meters) or strictly above are indicated in figure 5. As expected, the sensitivity with K_z is higher at 30m than above.

The segregation effect has been parameterized by multiplying the kinetic rate of the formation of NO_2 from O_3 and NO. The impact is plotted in figure 6. Notice that the real value is highly uncertain (the segregation effect is usually not described by comprehensive 3D Chemistry-Transport-Models).

The sensitivity with respect to emissions is illustrated in figure 7. A key result is the strong impact of NO emissions.

The dry deposition velocity of O_3 is also a key parameter as compared to the other deposition velocities (figure 8). The parameterization for cloud attenuation of photolysis is also not well known, due to the difficult diagnosis of clouds, hence, the sensitivity plotted in figure 9 (that has a quite low value as compared to other ones) may be underestimated.

To summarize, these preliminary results emphasize the impact of the boundary conditions for ozone and the impact of NO_x emissions.

Notice that the cost function J does not systematically admit a local minimum in the range $[0.5, 1.5]$ for α.

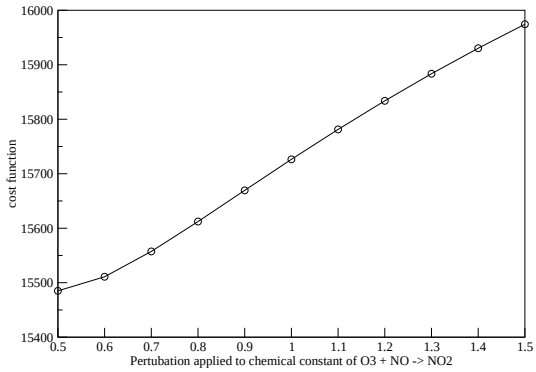

Figure 6. Dependence of the cost function on the kinetic rate of $O_3 + NO \rightarrow NO_2$.

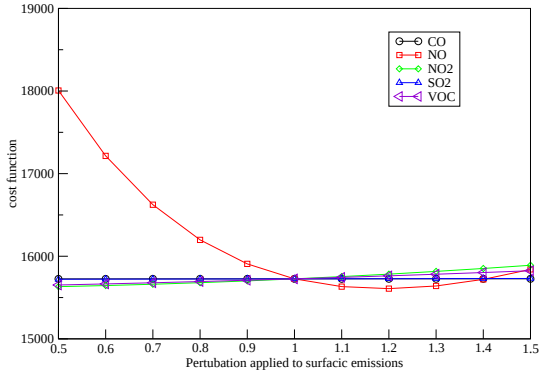

Figure 7. Dependence of the cost function on emissions.

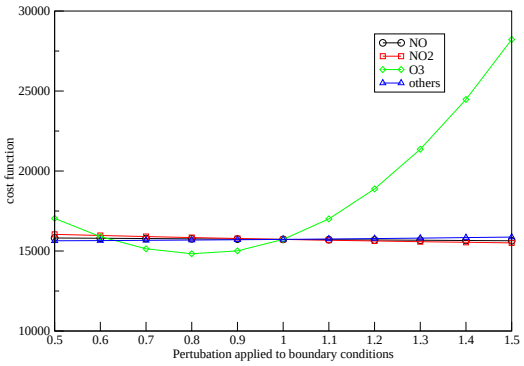

Figure 4. Dependence of the cost function on boundary conditions.

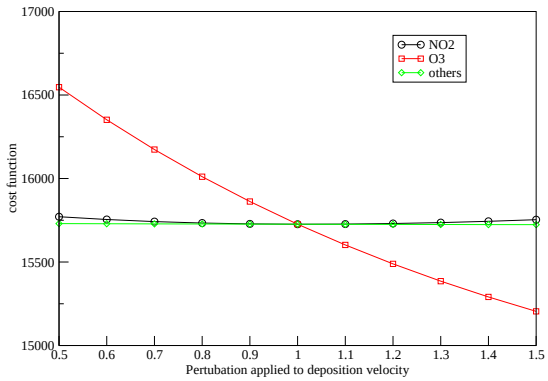

Figure 8. Dependence of the cost function on dry deposition velocities.

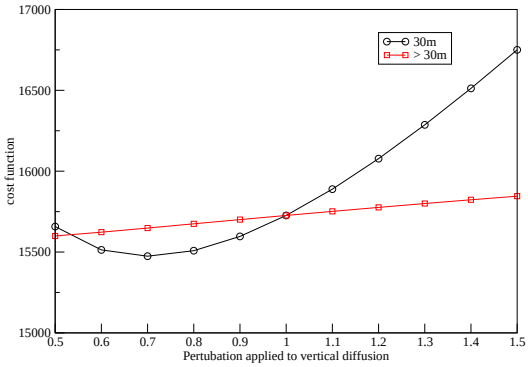

Figure 5. Dependence of the cost function on K_z.

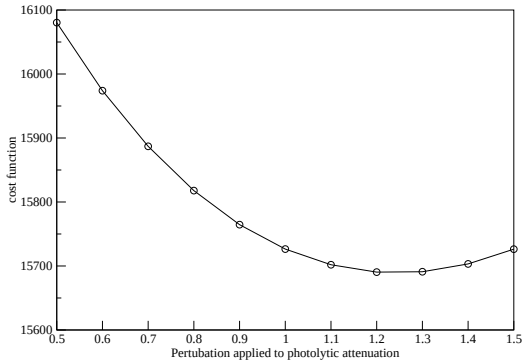

Figure 9. Dependence of the cost function on cloud attenuation.

3.2. Time and Space Impact of NO_x Emissions

A key issue for inverse modeling is to reduce the dimension of the control space (that is to say the number of degrees of freedom to be optimized). The emission data is given by 3D fields (two dimensions for surface emissions and one dimension for time) for every emitted species, which gives a very large number of parameters. According to the previous tests, the emission data that have the greatest impact by far are the NO_x emissions. Hence the control space should be reduced by first discarding all other emitted species.

To further reduce the control space, we now investigate the space and time impact of given NO_x emissions. We proceed in the following way:

1. for the spatial impact, the NO_x emissions of the grid cell $(15, 10)$ (arbitrarily chosen) are perturbed by $+30\%$ at each time step. The differences in the daily averages are then computed for species O_3 (figure 12), NO (figure 10) and NO_2 (figure 11).

2. for the time impact, a perturbation of $+30\%$ is applied at 0300 UT to NO_x emissions in all grid cells. We then compute the time evolution (on a hourly basis) of the differences in the spatial averages.

The impact of a perturbation in NO_x emissions is highly local in space. The largest difference for the three species output (NO, NO_2, O_3) is located in the grid cell where the emission occurs. In the vicinity of this cell, the difference is reduced by a factor of 20 for NO and by a factor of 6 for NO_2 or O_3. A few kilometers further, the emissions of NO_x have no more impact.

The figure 13 illustrates how long a perturbation in NO_x emissions has an impact on the monitored concentrations. After 3 hours, the impact may be neglected. This means that an observation may give some information for emissions in the three previous hours. This may be partially related to the residence time in the domain.

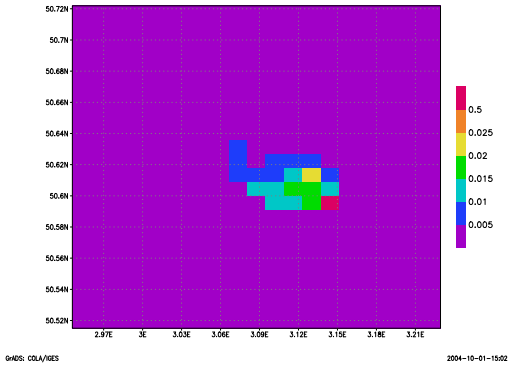

Figure 10. Map of the difference in averaged concentrations for NO due to a perturbation in NO_x emission in a given cell.

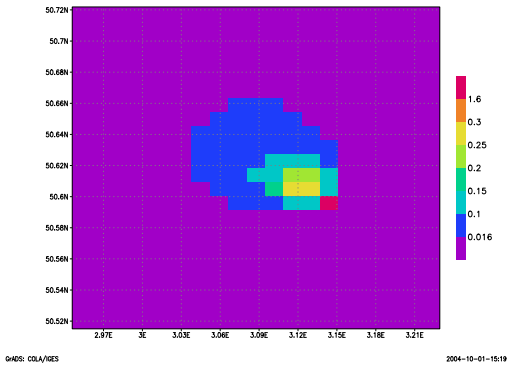

Figure 11. Map of the difference in averaged concentrations for NO_2 due to a perturbation in NO_x emission in a given cell.

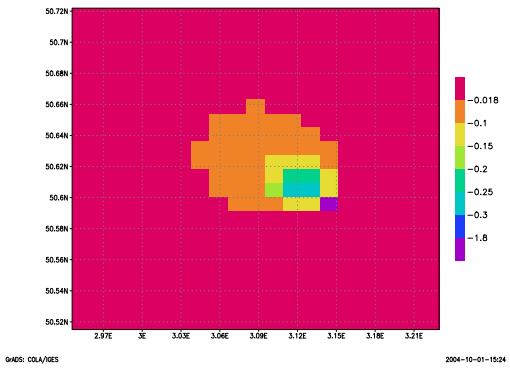

Figure 12. Map of the difference in averaged concentrations for O_3 due to a perturbation in NO_x emission in a given cell.

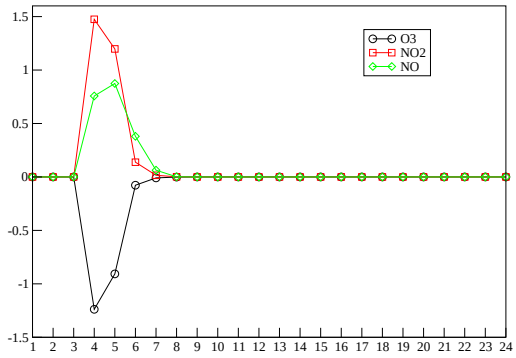

Figure 13. Time evolution of the difference in averaged concentrations for O_3, NO and NO_2 due to a perturbation in NO_x emissions at 0300 UT.

4. Twin Experiments for Inverse Modeling of Emissions

4.1. Some Notations

4.1.1. Control Variables

An accurate emission inventory has been made over Lille, in the framework of the French Program PREDIT (program of research, experimentation and innovation in land transport) in order to evaluate the health impact of emissions (`http://www.certu.fr/doc/env/echange/predit/predit.htm`). The spatial distribution of emissions is therefore assumed to be fairly accurate. The situation is quite different for the time distribution, which is given by monthly, daily and hourly coefficients. The coefficients are derived from average situations.

We have thus chosen, as control parameters, the time distribution of the most sensitive emitted species, namely NO_x. The emissions are parameterized with the form:

$$E_{\mathrm{NO_x}}(x,t) = \alpha(t)\tilde{E}_{\mathrm{NO_x}}(x,t) \tag{4}$$

where $\tilde{E}_{\mathrm{NO_x}}(x,t)$ are the emissions given by the emission inventory and $\alpha(t)$ are hourly coefficients applied to emissions over the whole domain. There are therefore 24 coefficients per day. In the reference case, $\alpha(t) = 1$ and we try to perform inverse modeling of $\alpha(t)$. We have chosen to focus on weekdays (that have different emissions as compared to weekends).

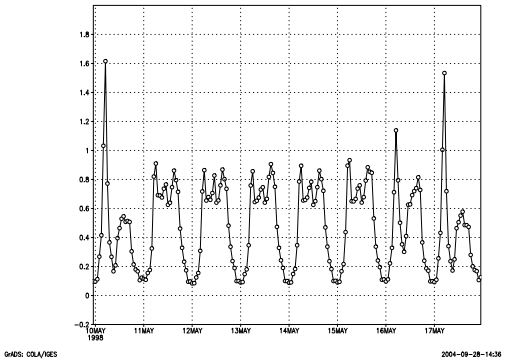

Figure 14. Time distribution for NO_x emissions over Lille from 10 May (Sunday) to 17 May (Sunday).

The choice of the control parameters is ruled by (at least) two criteria:

1. the impact on the cost function has to be large enough (see section 3);

2. they are supposed to be valuable for other days. Notice the time distributions for the week days are similar (figure 14).

The purpose is then to get a time distribution for traffic emissions of NO_x for the Monday-Friday period.

4.1.2. Choice of the Cost Function

Inverse modeling requires the specification of error statistics for three kinds of data: model outputs, measurements and control parameters (the so-called background terms). It is usually assumed that the model is perfect and we follow the same assumption. We have also chosen not to include a background term in the cost function: mathematically speaking, this term may be viewed as a penalty term or a regularization needed by ill-posed problems. In our case, the number of observations is larger that the number of control parameters so that this term is not required. Moreover we want to allow a strong sensitivity of the results with respect to the background term. The observational error is assumed to be constant over all the monitoring stations.

On the basis of these assumptions, the cost function is:

$$J(\alpha) = \sum_i \left((Hf(\alpha\tilde{E}_{NO_x}))_i - obs_i \right)^2 \tag{5}$$

where i labels the time, space and chemical "position" of the observations. We have only written the dependence of the model output with respect to NO_x emissions. The model is taken as a strong constraint (no model error), a constant weight is given to all observations and no background term is included.

We now want to minimize $J(\alpha)$ with respect to α.

4.1.3. Numerics and CPU Performance

This minimization problem may be solved by many algorithms. We have chosen to use the iterative algorithm BFGS [*Byrd et al.*, 1995] which belongs to the family of gradient algorithms. This requires to have at our disposal the gradient $\nabla_\alpha J$.

As previously mentioned, Polair3D has been built in order to have an adjoint mode easily available through automatic differentiation [*Mallet and Sportisse*, 2004]. Polair3D is written in Fortran 77 and may therefore be automatically differentiated, as it is, by O∂yssée [developed at Inria, *Faure and Papegay*, 1998]. To reduce the computational costs, only the differentiated LU factorization and solver for the chemistry are replaced. The main constraint in the process is to properly make the calls to the differentiated code and to check the validity of the adjoint with finite differences comparisons (the so-called Taylor tests).

The ratio of the CPU time needed for the adjoint computation (21 minutes per day with a 3 GHz processor) to the CPU time needed for the forward computation (about 3 minutes) is approximatively 7. This ratio is not optimal but this is not an issue in this case.

Each iteration of the procedure needs many evaluations of the cost function and of the gradient. For instance, 50 iterations of BFGS require 20 hours of CPU time.

No stopping criterion was used. As we were not in an operational context, we have chosen to let the iterative algorithm go to convergence. Notice that the convergence is mainly reached during the first steps of the process and that it is possible to limit the number of iterations without lowering the accuracy of the results (see below).

4.2. Twin Experiments without Perturbations

The first purpose is to validate the numerical algorithms (gradient computation, minimization) on the basis of twin

experiments without perturbations. It consists in using numerical data on observational data. These numerical observations are generated on the basis of "true" values for the control parameters: α^t (t stands for true). We then try to recover these values starting from a first guess in the minimizing algorithm α^b (b stands for background).

In practice we use $\alpha_i^t = 1$ and we have chosen to have 30% of underestimation for the first guess: $\alpha_i^b = 0.7$

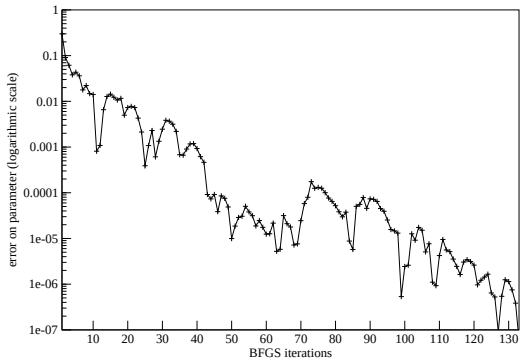

Figure 15. Error for α_8 as a function of the iterations of BFGS.

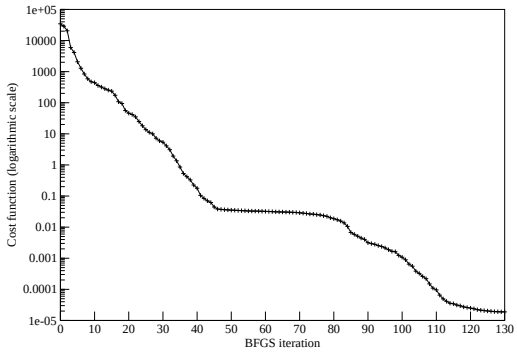

Figure 16. Evolution of the cost function with respect to the iterations of BFGS.

We have plotted the evolution of the cost function with respect to the iterations of BFGS in the figure 16. The true parameters are recovered after optimization up to round-off errors (see for instance the figure 15 for the parameter α_8).

4.3. Twin Experiments with Perturbations

The former experiment has been done without any perturbations and has proven the validity of the numerical tools. Before applying the approach to a real case it is however necessary to evaluate the ability of the system to deal with perturbations. The underlying issue is to have an estimation of the quality of the results.

We have chosen to perform two experiments:

1. the first experiment is related to observational errors: the numerical observations are therefore perturbed;

2. the second experiment is related to model errors: some input parameters of the model (see below for the details) are perturbed in order to generate the numerical observational. It is an easy way to generate model errors.

4.3.1. The impact of Observational Errors

The observational error has in practice two components: a first component is related to the errors made in the measurement process; a second component is the so-called error of representativeness and is related to the mismatch between the observation resolution (for instance a point measurement) and the model resolution (for instance a cell of 1 kilometer by 1 kilometer). The error of representativeness is probably the most important part, for chemical data, due to the heterogeneity of the chemical species concentrations in the vicinity of sources at small scales.

We have chosen to apply a Gaussian perturbation to the model outputs in order to parameterize the observational error. This error is assumed:

1. to be uncorrelated between chemical species and between different spatial locations;

2. to be correlated in time in order to avoid unrealistic fluctuations of observational errors;

3. to be larger for NO than for NO_2 and O_3 (with a ratio of 2) due to the local nature of NO. Notice this is a value chosen arbitrarily for these twin experiments.

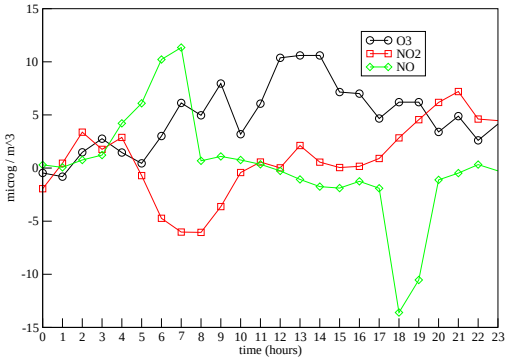

Figure 17. Observational errors generated for a given monitoring station in Lille for 11 May.

A typical example of such an observational error generated by this approach is plotted in figure 17. The impact is low for the optimized set of parameters (see figure 18).

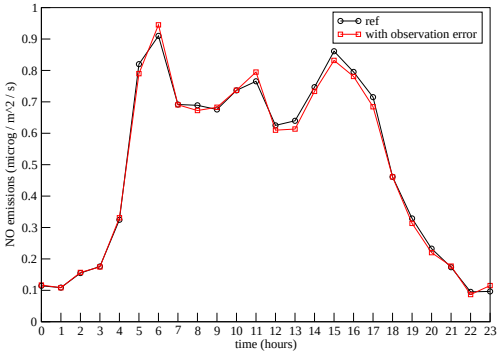

Figure 18. NO emissions: reference values and values obtained with observational error randomly generated.

4.3.2. The Impact of Model Errors

We can distinguish three kinds of model errors:

1. The errors related to forcing fields: meteorological data, dry deposition parameterizations, boundary conditions, etc. This corresponds in practice to all input parameters that may be uncertain but that are not optimized.

2. The errors related to the model itself: for instance, the segregation effects are neglected and the kinetic rates are used as if the fields were well mixed.

3. The errors related to numerical algorithms: for instance, the numerical scheme used for advection may induce large numerical diffusion.

These model errors lead to discrepancies between the model outputs and the observational data. A minimization of the cost function with respect to the emission parameters may be only a fit to observational data while the true reason for a large value of the cost function may be related to a large model error. For instance, a reduction of NO_x emissions may have a similar effect as an increase in the dry deposition velocities of NO_x.

In practice, we have perturbed independently each input parameter except the NO_x emissions. For each input field, an homogeneous value has been applied at each time. Notice that we have assumed that the model error is unbiased.

It is therefore important to assess the impact of model errors on the quality of the inverse modeling. We have chosen to parameterize the model errors by applying a Gaussian perturbation to the input parameters listed in section 3, except for the emissions of NO_x. We have assumed that the uncertainties follow a Gaussian law with a variance of 50%, which is coherent with the values given in *Hanna et al.* [2001].

The minimization of the cost function leads to a reduction of 85%. The new emission profile is plotted in figure 19. It is noteworthy that the period 0600-0700 UT in the morning is highly sensitive to model errors.

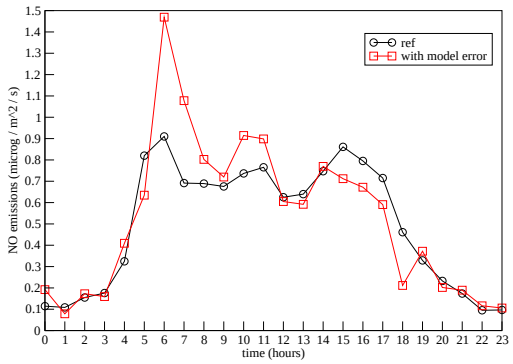

Figure 19. Optimized emissions of NO: reference values and values obtained with a model error.

4.4. Brief Summary of the Twin Experiments

The preliminary case with the numerical observations has led to the following conclusions:

1. the numerical algorithms (including the adjoint model) are validated.

2. the quality of the optimized parameters is not strongly lowered by observational errors and model errors.

The next section describes the application to the real case over Lille.

5. Application to a Real Case

The observations are now provided by the stations of the monitoring network AREMA. The forward model is effective for modeling O_3 and NO_2 but is not able to give an accurate forecast of NO. We have decided then to perform inverse modeling with and without the observations of NO. In a first experiment, we have at our disposal 4 observations of O_3 and 10 observations of NO_2 per hour.

Our approach is summarized in the following way:

1. the time distribution of $\alpha(t)$ is optimized during a learning period (typically one week).

2. we check the improvements of the emission inventory by using the optimized set of parameters during a verification period (typically a few weeks after the learning period).

5.1. Inverse Modeling from 11 May to 15 May

We have chosen the week from 11 May to 15 May as a learning period. Two kinds of experiments have been performed in order to estimate the daily variability of the optimized distribution:

1. in a first approach, each day is independently used as a learning period (5 learning periods), which leads to 5 sets of optimized parameters.

2. in a second approach, the week (actually 5 days, the weekend being excluded) is used as a learning period as a whole, which leads to a unique set of optimized parameters.

In practice we have added a simulation period of 6 hours before the beginning of the periods in order to take into account the model spin-up and to lower the influence of the initial conditions.

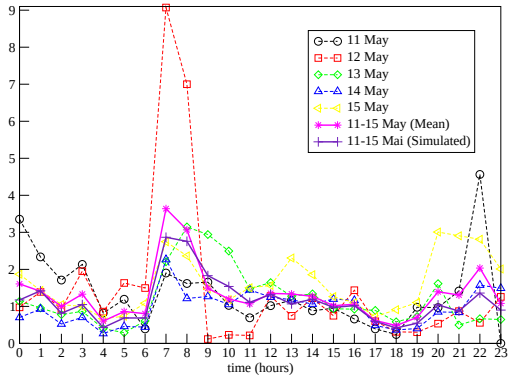

Figure 20. Daily distribution of the optimized parameters α for the learning periods from 11 May to 15 May.

The optimized parameters α are plotted in figure 20 for the different learning periods defined above. For the first experiment (learning periods of one day), the convergence of the minimization algorithm (with a strong requirement for the stopping test) is obtained after a number of iterations ranging from 59 to 128. The reduction of the cost function ranges from 25% to 66% (see table 1). The case of the global learning period (second case) requires 42 iterations and leads to a reduction in the cost function of 20%.

The runs for the one-day learning periods illustrate the rather high variability of the optimized parameters. However, there are some similar features:

1. the coefficients corresponding to hours 7 and 8 are greater than 1.

2. the coefficients corresponding to hours 17, 18 and 19 are lower than 1.

These two periods are highly sensitive since they are related to the emissions peaks but also to key evolutions of other processes (photolysis and vertical mixing).

12 May is not well modeled: the initial cost function has high values. We suspect then that other parameters have bad values and that the model error (as defined above) is large.

We also observe an overestimation of α in the night from Sunday to Monday and in the night from Friday to Saturday. One possible realistic reason could be the overestimation of traffic-jams related to weekends.

Table 1. Values of the cost function before and after optimizations; performance of the convergence of the BFGS algorithm for the different learning periods.

	11 May	12 May	13 May	14 May	15 May	11-15 May
Initial cost function (10^{-5})	5.38	9.50	7.65	3.85	2.96	16.
Optimized cost function (10^{-5})	2.55	3.03	4.40	2.91	2.06	12.8
Iterations of BFGS	101	128	59	64	55	42

The plots labelled by "11-15 May (Mean)" are related to the average of the 5 optimized sets of parameters obtained with a one-day learning period. Notice that the distribution is similar to the distribution obtained with a five-day learning period, labelled by "11-15 May (Simulated)".

The time distribution obtained for the emissions is then plotted in figure 21. One key remark is of course the lack of symmetry between the morning and the end of the afternoon. There are many possible reasons for that: this could be due to model errors (for instance the extension of the mixing layer in the morning); another reason that could explain these features is related to different regimes of emissions (cold emissions in the morning) that are perhaps not well represented by the emission inventory.

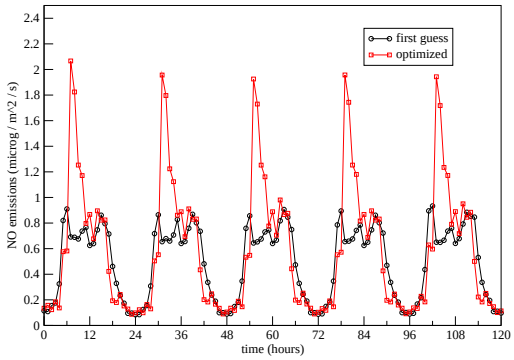

Figure 21. Daily distribution of NO emission for the period 11-15 May. Reference and optimized parameters.

5.2. Verification

We have checked the improvement of the emission inventory by applying the optimized time distribution to other weeks than the learning week. The Root Mean Square Errors (RMS) and the correlations computed for the learning week and the two weeks after are given in Table 2. The forecast skills are improved with the exception of the RMS for NO during week 25-30 May.

5.3. Use of other Learning Periods

The variability of the optimized parameters with respect to the learning period is a way to investigate the robustness of the approach. As one can assume that there is no key reason justifying that the time distribution has a drastic evolution from one week to another, this is more or less equivalent to investigating the sensitivity to meteorological conditions.

The optimized parameters for three different learning periods (week 11-15 May and the two weeks after) are plotted in figure 22. We observe similar behaviors for the three learning periods: high values in the morning and low values at the end of the afternoon. The fact that the results are not highly sensitive to the learning period is an indication that the approach is robust.

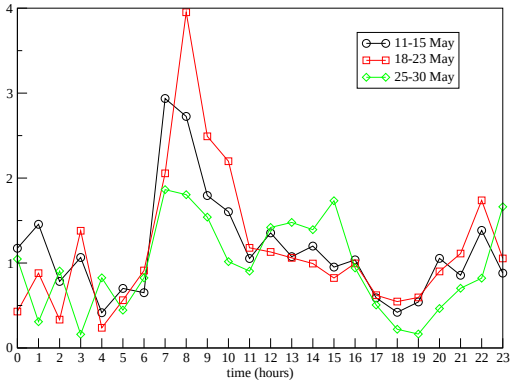

Figure 22. Optimized time distributions for three different learning periods. The "reference curve" stands for a first guess equal to 1.

5.4. Second-Order Sensitivity

The results of the inverse modeling procedure depend on many parameters that may be uncertain: the meteorological conditions, the parameterizations, other emissions, etc. They also depend on some parameters referred to as "assimilation parameters": for instance, the first guess, the covariances matrices (if any, which is not the case in this study) or the observational operator.

Mathematically speaking, we can write the cost function J with a more general form as $J(\alpha, k_p, k_a)$ where k_p stands for the physical parameters supposed to be known (but uncertain) and k_a for the assimilation parameters.

The optimized values $\alpha^\star$ are given by:

$$\nabla_\alpha J(\alpha, k_p, k_a) = 0 \tag{6}$$

which defines a function $\alpha^\star(k_p, k_a)$. The second-order sensitivity deals with the sensitivity of these optimized parameters with respect to k_p (robustness with respect to physical parameters) and k_a (impact of the monitoring network, typically).

A comprehensive way to assess this sensitivity is to compute the partial derivatives of $\alpha^\star$ with respect to k_p and k_a. This may be done with the Hessian matrix of J [*Le Dimet et al.*, 2002].

It is of course a huge task to develop this Hessian matrix when J is related to 3D comprehensive models, such as Polair3D. Even if this second-order mode is available [*Sportisse and Quélo*, 2003], we have chosen to use another simpler approach. We have chosen to compute the optimized results $\alpha^\star$ with perturbed values of some assimilation parameters and of physical parameters. More precisely, we have restricted the study to the impact of the first guess, the impact of NO observations and the impact of K_z (which describes the mixing layer).

5.4.1. Impact of the First Guess

Table 2. RMS errors and correlations (in brackets) before (reference) and after optimization (optimized) of the time distribution α during week 11-15 May.

	11-15 May		18-13 May		25-30 May	
	Reference	Optimized	Reference	Optimized	Reference	Optimized
O_3	32.7 (0.83)	29.6 (0.87)	19.9 (0.82)	18.9 (0.85)	20.8 (0.61)	18.1 (0.69)
NO_2	29.4 (0.36)	25.5 (0.52)	19.9 (0.49)	18.4 (0.59)	19.4 (0.24)	17.4 (0.29)
NO	27.4 (0.61)	26.6 (0.66)	19.7 (0.49)	17.7 (0.66)	20.3 (0.38)	21.5 (0.40)

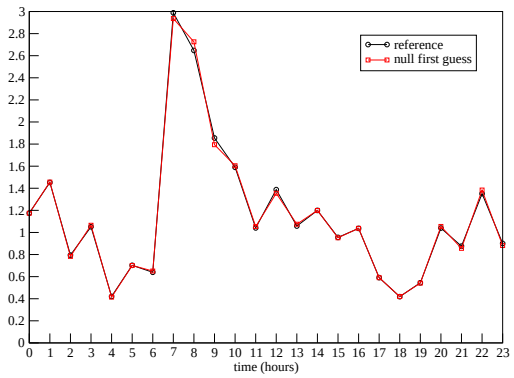

Figure 23. Optimized time distribution for two sets of first guesses for α (1 or 0).

A first experiment has been performed in order to check the sensitivity with respect to some assimilation parameters, such as the first guess for α. As said before, we have not used a penalty term for the cost function (usually referred to as a background term) because the problem is well posed and does not require a regularization technique. What we call here a "first guess" is the choice of the reference values for α, that is, the initial values used for starting the minimization algorithm.

We have compared two choices for first guess: the first guess is either 1 (that is to say we start from the time distribution given by the emission inventory) or is 0 (lack of information for the time distribution of emissions at the beginning).

A key result is that we obtain the same distribution after optimization (see figure 23). The number of iterations required for convergence is however larger for the second case as the starting point is farther from the optimal values (27 iterations against 22 in the first case). Mathematically speaking, this proves that the cost function has a strongly convex behavior in a large domain around the optimal set of parameters.

This result proves that the optimized time distribution may be recovered without any use of a priori information given by the emission inventory.

5.4.2. Sensitivity with Respect to the Observed Species

We have chosen up to now not to take into account the observations of NO. We investigate in this subsection the impact of these observations. We have still assumed that the observational errors for all species are the same ones (even if this is probably not the case: NO has for instance a larger error of representativeness due to its local nature). The results are plotted in figure 24. We recover a similar qualitative behavior as before, even if the overestimation in the morning is increased. This is consistent with the strong underestimation of NO concentrations by the model in the morning.

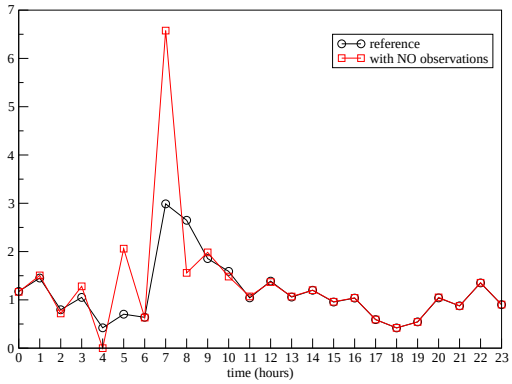

Figure 24. Optimized distribution by adding observations of NO. The reference curve stands for the optimized α without NO observations.

5.4.3. Sensitivity with Respect to Vertical Mixing

We have already mentioned the sensitivity at the beginning of the morning, which corresponds to the transition from the nocturnal stable boundary layer to a mixed layer. It is however well recognized that the parameterization of K_z has a key impact on model outputs. Another reason could be the impact of sunset through photolysis but our tests (not reported here) do not indicate a strong sensitivity.

In order to assess the sensitivity of the results with respect to K_z, we have artificially decided to shift forward the time distribution of K_z by one hour. The start of the extension of the mixed layer then occurs one hour after the reference case.

The optimized parameters are plotted in figure 25 for one day (11 May). As expected, the main differences are obtained during the transition periods in the morning and in the afternoon. The time distribution is however not highly modified, which confirms the robustness of the result.

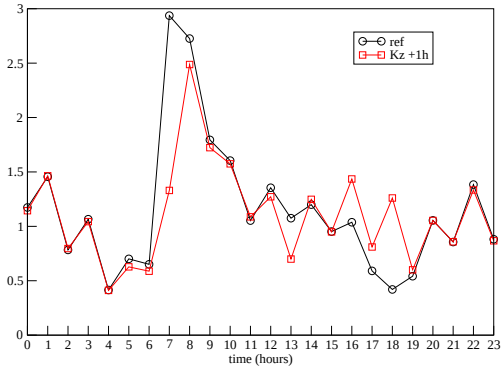

Figure 25. Optimized distribution for 11 May: with the reference K_z and with a shifted K_z.

Conclusions and Future Works

We have applied a variational approach in order to perform inverse modeling of emissions at regional scale, over Lille in northern France. After a sensitivity analysis, we have chosen to optimize the time distribution of NO_x on the basis of observations of O_3, NO_2 and NO.

Twin experiments have proven the validity of the numerical models (especially the adjoint model of our Chemistry-Transport-Model, Polair3D, obtained by automatic differentiation). We have also tested the impact of observational errors and of model errors in numerical tests.

The application to one week of May 1998 has led to an optimized set of parameters. A verification test (by applying the optimized distribution to the next two weeks) has confirmed the improvement of the forecast skills.

A brute-force second-order sensitivity analysis has also been performed in order to check the robustness of the optimized parameters with respect to other uncertain parameters (first guess, meteorological conditions, K_z). The results have proven that the optimized time distribution is robust.

Future work will be devoted to the application at continental scale (with a focus on spatial distribution rather than time distribution) with a focus on second-order sensitivity. Another key point could also be to take into account model errors in the inverse modeling process. Many approaches are under investigation, ranging from combined inverse modeling of other parameters than emissions to weak formulations of the variational problem or Monte Carlo simulations of the inverse modeling procedure.

Acknowledgments. We thank the PREDIT program and the CETE institute for providing us the emission inventory for road traffic used in this paper. We also thank Rémy Lagache for fruitful discussions about emission inventories.

References

Bergamaschi, P., R. Hein, M. Heinmann, and P. Crutzen, Inverse modeling of the global co cycle. 1. inversion of co mixing ratio, *J.Geophys.Res.*, *105*(D2), 1909:1927, 2000.

Bousquet, P., P. Ciais, P. Peylin, M. Ramonet, and P. Monfray, Inverse modeling of annual atmospheric CO2 sources and sinks 1. method and control inversion, *J. of Geophys. Res.*, *104*(D21), 26,161–26,178, 1999.

Boutahar, J., S. Lacour, V. Mallet, D. Quélo, Y. Roustan, and B. Sportisse, Development and validation of a fully modular platform for numerical modelling of air pollution: POLAIR, *Int. J. Environment and Pollution*, *22*(1/2), 17–28, 2004.

Byrd, H., P. Lu, J. Nocedal, and C. Zhu, A limited memory algorithm for bound constrained optimization, *SIAM J. Scientific Computing*, *16*(5), 1190–1208, 1995.

Chang, J., R. Brost, I. Isaken, S. Madronich, P. Middleton, W. Stockwell, and C. Walcek, A three-dimensional eulerian acid deposition model: physical concepts and formulation, *J. Geophys. Res.*, *92*(D12), 14,681–14,700, 1987.

Chang, M. E., D. E. Hartley, C. Cardelino, D. Haas-Laursen, and W.-L. Chang, On using inverse methods for resolving emissions with large spatial inhomogeneities, *J. Geophys. Res.*, *102*(D13), 16,023–16,036, 1997.

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, and E. Ebel, 4d variational data assimilation with an adjoint air quality model for emission analysis, *Environ. Mod. Softw.*, *15*, 539–548, 2000.

Faure, C., and Y. Papegay, Odyssée user's guide. version 1.7, *Tech. Rep. RT-0224*, INRIA, 1998.

Hanna, S. R., Z. Lu, H. C. Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, and D. A. Hansen, Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the july 1995 otag domain, *Atmos. Env.*, *35*, 2001.

Horowitz, L. W., et al., A global simulation of tropospheric ozone and related tracers: description and evaluation of MOZART, version 2, *J. Geophys. Res.*, *108*(D24), 2003.

Hourdin, F., and J.-P. Issartel, Sub-surface nuclear tests monitoring through the ctbt 133xe network, *Geophys. Res. Lett.*, *27*, 2245–2248, 2000.

Kaminski, T., On the benefit of the adjoint technique for inversion of the atmospheric transport employing carbon dioxide as an example of a passive tracer, Ph.D. thesis, Un. Hamburg, 1998.

Le Dimet, F.-X., I. Navon, and D. Daescu, Second-order information in data assimilation, *Month. Weat. Rev*, *130*, 629–648, 2002.

Louis, J.-F., A parametric model of vertical eddy fluxes in the atmosphere, *Boundary-Layer Meteorology*, *17*, 187–202, 1979.

Madronich, S., Photodissociation in the atmosphere: 1. actinic flux and the effects of ground reflections and clouds, *J. Geophys. Res.*, *92*(D8), 9,740–9,752, 1987.

Mallet, V., and B. Sportisse, 3-D chemistry-transport model Polair: numerical issues, validation and automatic-differentiation strategy, *Atmos. Chem. Phys. Discuss.*, *4*, 1,371–1,392, 2004.

Mallet, V., and B. Sportisse, Data processing and parameterizations in atmospheric chemistry and physics: the AtmoData library, *Submitted to Atmospheric Environment*, 2005.

Mallet, V., D. Quélo, and B. Sportisse, Software architecture of an ideal modeling platform in air quality – A first step: Polyphemus, *Submitted to Atmospheric Environment*, 2005.

Mendoza-Dominguez, A., and A. Russel, Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling, *Atmos.Environ.*, *35*, 2879–2894, 2001.

Menut, L., Adjoint modelling for atmospheric pollution processes sensitivity at regional scale during the ESQUIF IOP2, *J. of Geophys. Res.*, *108*(D17), 2003.

Middleton, P., W. R. Stockwell, and W. P. L. Carter, Aggregation and analysis of volatile organic compound emissions for regional modeling, *Atmos. Env.*, *24A*(5), 1,107–1,133, 1990.

Quélo, D., Simulation numérique et assimilation de données variationnelle pour la dispersion atmosphérique de polluants, Ph.D. thesis, École Nationale des Ponts et Chaussées, 2004.

Sandu, A., F. Potra, G. Carmichael, and V. Damian, Efficient implementation of fully implicit methods for atmospheric chemical kinetics, *J.Comp.Phys.*, *129*, 101–110, 1996.

Segers, A., Data assimilation in atmospheric chemistry models using kalman filtering, Ph.D. thesis, TU Delft, 2002.

Simpson, D., et al., Inventorying emissions from nature in Europe, *J. Geophys. Res.*, *104*(D7), 8,113–8,152, 1999.

Sportisse, B., and D. Quélo, Data assimilation and inverse modeling of atmospheric chemistry, *Proc. of Indian National Science Academy. Part A Physical Sciences*, *69*, 2003.

Stockwell, W. R., F. Kirchner, M. Kuhn, and S. Seefeld, A new mechanism for regional atmospheric chemistry modeling, *J. Geophys. Res.*, *102*(D22), 25,847–25,879, 1997.

van Loon, M., P. Builtjes, and A. Segers, Data assimilation of ozone in the atmospheric transport chemistry model LOTOS, *Env. Mod. and Soft.*, *15*, 603–609, 2000.

Verwer, J. G., W. H. Hundsdorfer, and J. G. Blom, Numerical time integration for air pollution models, *Tech. rep.*, CWI, 1998.

Wesely, M. L., Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmos. Env.*, *23*, 1293–1304, 1989.

D. Quélo, CEREA, École Nationale des Ponts et Chaussées, 6-8 Avenue Blaise Pascal, 77455 Champs sur Marne, France. (quelo@cerea.enpc.fr)

V. Mallet, CEREA, École Nationale des Ponts et Chaussées, 6-8 Avenue Blaise Pascal, 77455 Champs sur Marne, France. (mallet@cerea.enpc.fr)

B. Sportisse, CEREA, École Nationale des Ponts et Chaussées, 6-8 Avenue Blaise Pascal, 77455 Champs sur Marne, France. (sportiss@cerea.enpc.fr)