Available online at www.sciencedirect.com
ScienceDirect
Computers and Mathematics with Applications 55 (2008) 2329–2337
An International Journal
computers & mathematics with applications
www.elsevier.com/locate/camwa

Air quality modeling: From deterministic to stochastic approaches

Vivien Mallet, Bruno Sportisse*

Université Paris-Est, CEREA (joint laboratory ENPC - EDF R&D), Marne la Vallée, France
INRIA, Paris-Rocquencourt Research Center, France

Abstract

The objective of this article is to investigate the topics related to uncertainties in air quality modeling. A first point is the evaluation of uncertainties for model outputs: Monte Carlo methods and sensitivity analysis are powerful methods for assessing the impact of uncertainties due to model inputs. A second point is devoted to ensemble modeling with multi-models approaches. According to the wide spread in the model outputs, using a unique model, tuned to a small set of observational data, is not relevant in this field. On the basis of ensemble simulations, improved forecasts are given by appropriate algorithms to combine the set of models. The results applied to air quality modeling at continental scale with the POLYPHEMUS system illustrate these methods. The first estimates of uncertainties in inverse modeling experiments are also proposed.

Keywords: Uncertainty analysis; Sensitivity analysis; Ensemble forecast; Data assimilation; Air quality

0. Introduction

Air quality models describe the atmospheric dispersion of trace species. The species may be inert species (for instance radionuclides or heavy metals) or reactive species (for instance gas-phase compounds that participate in photochemistry, such as ozone). Particle matter (aerosols) can also be taken into account. These models are applied from local scale (in the vicinity of emissions) to regional and continental scales.

The applications of such models range from process studies (the understanding of key chemical and physical phenomena) to operational forecast (similar to numerical weather prediction) and impact or scenario studies (What would happen if emissions were modified?).

One usually refers to the so-called chemistry-transport models (referred to as CTM in what follows) at regional and continental scales. These models simulate the time evolution of 3D fields of species concentrations. If c_i stands for the concentration of a gas-phase species labelled by i, the time evolution is governed by a reaction-diffusion-advection partial differential equation:

$$\frac{\partial c_i}{\partial t} + \operatorname{div}\,(V c_i) = \operatorname{div}\left(\rho K \nabla \frac{c_i}{\rho}\right) + \chi_i(c, x, t) - \Lambda_i(x, t) c_i + S_i(x, t). \tag{1}$$

* Corresponding address: ENPC - CEREA, 6-8 avenue Blaise Pascal, Champs-sur-Marne, 77455 Marne-la-Vallée Cedex 2, France. Fax: +33 1 64 15 21 70.
E-mail addresses: mallet@cerea.enpc.fr (V. Mallet), sportiss@cerea.enpc.fr (B. Sportisse).
URLs: http://cerea.enpc.fr/, http://www-rocq.inria.fr/clime/.

0898-1221/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2007.11.004

V is the wind velocity and ρ is the air density: they are computed by meteorological models. K is the eddy coefficient matrix and Λ_i is a scavenging coefficient (for instance for below-cloud and in-cloud scavenging for soluble species): they are computed by parameterizations as functions of the meteorological fields. S_i is the point (volume) source terms, given by emission inventories. χ_i is the chemical source term that describes the chemical reactions. It is a nonlinear term for reactive species. For species bound to aerosols, this term has a complicated form (this is an advection-integro-differential term, the size of the particle also being taken into account). Notice that for reactive species, these PDEs (Partial Differential Equation) are coupled through the chemical terms.

Initial conditions and boundary conditions have to be added to this set of PDEs. For instance, at ground:

$$-K\nabla c_i \cdot n = E_i(x,t) - v_i^{\text{dep}}(x,t)c_i$$

with n the vertical unitary vector (upward oriented), E_i the surface emissions and v_i^{dep} the dry deposition velocity, computed as a parameterization of meteorological fields and the land use cover.

Such models may describe up to hundreds of species (for tropospheric photochemistry and aerosols).

These models are characterized by many uncertainties:

- in the model inputs: the emission inventories are not very accurate (for spatial and time distribution, for chemical speciation), the meteorological fields are also uncertain, etc.
- in the parameterizations: a key point for atmospheric modeling is the need for subgrid parameterizations in order to describe the effects of unresolved processes that occur at scales (the physical scale) much smaller than the resolved scale (the numerical scale). At continental scale, the horizontal dimension of one grid cell is typically 50 km while many processes occur at scales ranging from micrometers (mass transfer between gas-phase species and particulate matter, liquid droplets) to meters (turbulent eddies).
- in the numerical algorithms: the computational burden may be particularly high for comprehensive CTM. In practice, a coarse resolution may be used (especially for aerosols).

The objective of this paper is to review a few approaches that can be used to assess and to take into account these uncertainties in model outputs. There is a gap between the small set of available observational data (especially for the chemical composition) and the amount of uncertain parameters. Models are usually tuned to show good forecast skills for specific targets (for instance: ozone daily peaks). An alternative approach is to use a set of models without tuning.

This article is structured as follows. In Section 1, brute-force approaches (Monte Carlo methods, sensitivity analysis with respect to continuous parameters, and model-to-model comparisons) are introduced in order to assess the propagation of uncertainties in the model. In Section 2, ensemble methods are presented: they are based on algorithms that combine the outputs of a set of models (typically 50) supposed to represent the spread due to uncertainties. The combination meets an optimal criterion based on model-to-data comparisons. Both sections are illustrated by air quality modeling at continental scales with the POLYPHEMUS system. Section 3 is briefly devoted to the issues of data assimilation. Section 4 sketches out the conclusions and future works.

1. Assessment of uncertainties

Following the introduction of this paper, one may distinguish the uncertainties related to input data, to parameterizations (that is to say to the physical model itself) and to numerical schemes (including the discretization problem).

1.1. Numerics

The impact of numerical approximations is measured through model-to-model comparisons. A reference simulation is perturbed by modifying numerical components: new numerical algorithms or discretization parameters are changed. For instance, a key algorithm is the advection solver which needs to describe sharp gradients in many applications (e.g., dispersion of radionuclides). The time step and the mesh resolution are examples of discretization parameters.

Table 1
Summary of agreement coefficients for main comparisons performed between numerical algorithms

Comparison	O_3	SO_2	NO	NO_2	HO
$\Delta t = 600$ s/$\Delta t = 1800$ s	54.7	80.1	28.7	59.4	43.3
Third-order advection/First-order advection	66.0	70.4	81.5	68.9	81.0
$K_h = 10\,000\ \mathrm{m}^2 \cdot \mathrm{s}^{-1}/K_h = 50\,000\ \mathrm{m}^2 \cdot \mathrm{s}^{-1}$	80.0	81.9	18.4	65.7	83.9
$\Delta t = 600$ s/$\Delta t = 30$ s	96.4	89.4	83.9	79.7	57.2
boundary condition with diffusion/ boundary conditions with chemistry	97.0	88.2	80.4	84.9	80.7
$K_h = 10\,000\ \mathrm{m}^2 \cdot \mathrm{s}^{-1}/K_h = 0\ \mathrm{m}^2 \cdot \mathrm{s}^{-1}$	97.9	84.2	90.7	85.4	94.5
Reference/Strang splitting	98.5	96.3	78.8	89.0	81.7
splitting sequence diffusion–chemistry/ splitting sequence chemistry–diffusion	99.6	98.9	10.4	61.4	68.3
Reference/adapted Rosenbrock	99.7	100.0	94.6	97.7	97.9
Advection at start of splitting sequence/ Advection at end of splitting sequence	99.9	93.1	86.4	88.0	74.8
Reference/internal splitting	99.9	98.8	89.5	91.0	73.5
Reference/modified Rosenbrock	100.0	100.0	83.4	97.1	67.0
Reference/source splitting	100.0	100.0	93.9	90.7	77.8

The comparisons are sorted according to the agreement coefficient for ozone.

One first defines an agreement coefficient between two simulations, that generate the fields A and B, with means $\overline{A}$ and $\overline{B}$, through:

$$\mathrm{Agr}(A, B) = \frac{\mathrm{card}\ \{(h, i, j, k)/|\Delta_{h,i,j,k}| < 5\%\}}{\mathrm{card}\{(h, i, j, k)\}}, \quad \text{with} \quad \Delta_{h,i,j,k} = \frac{A_{h,i,j,k} - B_{h,i,j,k}}{\frac{1}{2}(\overline{A} + \overline{B})} \tag{2}$$

where h, i, j and k label time and directions x, y and z, respectively.

If a change in the reference configuration leads to an agreement coefficient much less than 100%, one considers that the associated numerical issue is a source of uncertainties. Notice that the agreement coefficient may also be defined for given chemical "targets", and the sensitivity to the numerical scheme strongly depends on the target: maximum concentration, mean concentration, chemical species. For instance, it is much harder to accurately simulate the concentrations of a highly reactive species like HO than to simulate concentrations of O_3. As a consequence, the agreement coefficients should be carefully analyzed.

A comprehensive set of tests has been made by investigating the impact of advection solvers, the operator splitting methods, the time and space discretization and the chemical solvers. All tests are performed with the full chemistry-transport model in realistic conditions, based on a well-studied simulation. The simulation takes place over Western Europe during a summer week. The photochemical mechanism RACM [1] is used. We refer to [2] for further details.

A synthesis of such tests is given in Table 1 for five chemical species: ozone O_3, sulfur dioxide SO_2, nitric oxide NO, nitrogen dioxide NO_2 and hydroxyl radical HO. The main conclusion is that the advection scheme plays a major role in the integration process, even for a simulation with a complex photochemical mechanism. The results are sensitive to the time step if it goes beyond a certain limit (about 600 s). A high horizontal diffusion has a noteworthy impact on all species.

1.2. Input data

The model is now viewed under the input/output form $y = F(\Psi)$, where y stands for the outputs, Ψ for the inputs supposed to be uncertain.

Monte Carlo simulations. The first approach is based on Monte Carlo simulations. On the basis of expert knowledge, most of the uncertain inputs may be described by lognormal probability density functions (PDF in the sequel). See Table 2 for estimated uncertainties for a simulation over Western Europe. A simulation with 800 Monte Carlo runs

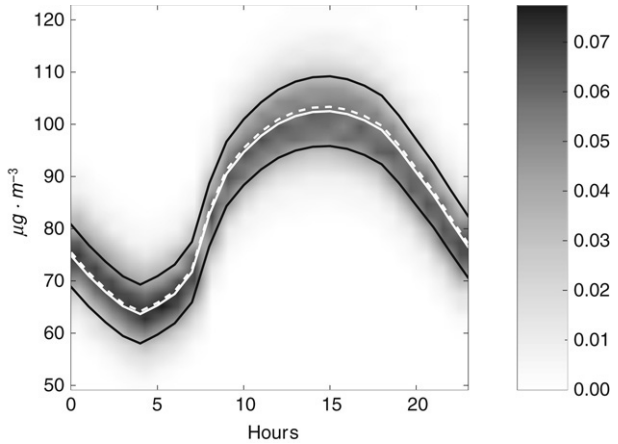

Fig. 1. Ozone mean daily profile. The mean is computed on the basis of ozone profiles in all ground cells over Europe and over one week in summer. The probability density is shown in the background (gray levels) with the expectation (continuous white line). The expectation plus and minus the standard deviation are represented by the black lines. The discontinuous white line is the profile of the reference simulation (that is, without perturbed inputs).

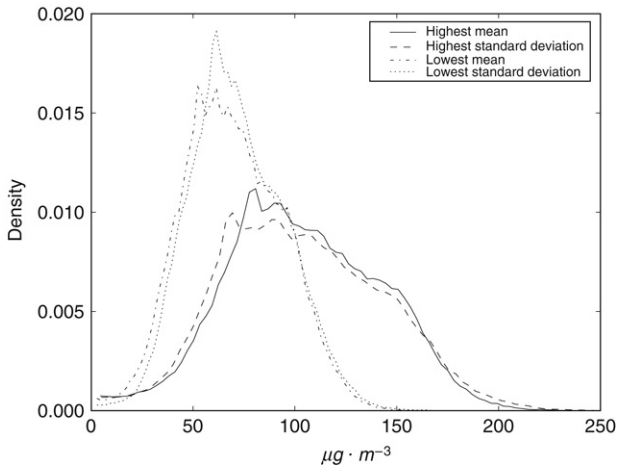

Fig. 2. Spatio-temporal distribution of ozone concentrations for extreme simulations in the Monte Carlo runs. These simulations show the highest mean concentration (103.4 μg m^{-3}), the highest standard deviation (40.3 μg m^{-3}), the lowest mean concentration (69.2 μg m^{-3}) and the lowest standard deviation (23.6 μg m^{-3}).

Table 2
Uncertainties associated with main input variables (except meteorological fields), based on [4,5]

Input data	Uncertainty (LN)(%)
Cloud attenuation	±30
Deposition velocities (O_3 and NO_2)	±30
Boundary conditions (O_3)	±20
Anthropogenic emissions	±50
Biogenic emissions	±100
Photolysis rates	±30

is presented in [3]. Typical results are illustrated with the approximated probability density function of ozone daily profiles (Fig. 1) and with the relative frequency distributions of simulations with extreme behaviors (Fig. 2).

The resulting computational burden is a key issue of such approaches. An alternative approach is based on reduced Monte Carlo simulations (for instance through variance reduction). Among other alternative methods, the so-called

chaos expansion is a powerful approach (for instance the DEMM method, [6]). The output is written as:

$$y = y_0 + y_1(\Psi) + y_2(\Psi) + \cdots + y_N(\Psi) \tag{3}$$

with the uncertain input Ψ that follows a given PDF. The sequence $(y_i)_i$ is a set of orthogonal polynomial functions (with respect to a given scalar product). One may refer to [7] for an application to air quality modeling over Europe. In this work, the studied uncertainties are related to NO_x, NH_3 and SO_2 emissions. The gain in CPU time may be up to 2 or 3 orders of magnitude.

Sensitivity analysis. Another way to assess the propagation of uncertainties is by computing the partial derivatives $\frac{\partial y}{\partial \Psi}$. A small perturbation $\delta\Psi$ in the inputs will lead to a perturbation $\frac{\partial y}{\partial \Psi}\delta\Psi$ in the outputs (after linearization). This sensitivity analysis is of course only limited to "local uncertainties".

Computing sensitivities may be a difficult task when the input parameter has a large dimension. The use of finite difference techniques may result in a too heavy computational burden (if n is the number of input parameters, $O(n)$ model runs are required). Moreover, the choice of small perturbations to compute finite differences may be a difficult task. An alternative approach is based on the so-called *linear tangent* model and *adjoint* model. When available, they can be a powerful technique for computing the sensitivities.

We refer for instance to [8] for the investigation of numerical issues, to [9] for an application to a diphasic model (gas phase and cloud droplet) and to [10] for the sensitivity of ozone levels with respect to emissions at continental scales.

1.3. Parameterizations

As pointed out in the introduction, a key issue for chemistry-transport models is the use of appropriate parameterizations for several processes. For a given input field (for instance the eddy vertical coefficient K_z), there exist different parameterizations. Moreover several data bases may also be used in the parameterizations (for instance, land use cover for dry deposition velocities). This results in different combinations of parameterizations and data. Each combination determines a model configuration. If the involved parameterizations and data bases are all *a priori* reliable, the results of any combination are equally likely. The spread in the model outputs from the set of model configurations is an estimate of *a priori* uncertainties. *A posteriori* uncertainties may also be computed by model-to-data comparisons, especially through Bayesian methods [11].

The POLYPHEMUS system [12] has been designed in such a way that many combinations of parameterizations and data bases may be used to define a simulation. A crucial point is of course the modularity, allowed by the use of software libraries (for instance the ATMODATA library for physical parameterizations), contrary to all-in-one models. A comprehensive analysis of *a priori* uncertainties due to parameterizations is presented in [13] for ozone modeling at continental scale. One conclusion is the wide spread in the model outputs in this multi-models approach.

An illustration is shown in Fig. 3 where profiles from 48 different models (mainly built with different combinations of parameterizations and data bases) is shown. Even if the profiles are concentrations averaged over many grid cells (about 2000) and several days (about 120), they show a wide spread. In our work, the uncertainty due to the model formulation is clearly higher than the uncertainty due to input data (except meteorological data). This shows that a simulation with a unique model is seldom sufficient for detailed studies.

2. Ensemble forecast

The assessment of uncertainties, whatever the source is, leads to the conclusion than model outputs from one single model (in one fixed configuration) are highly doubtful. A single model can be well adjusted (tuned) to have good forecast skills for a small set of targets (usually ozone peaks) but the number of observational data (especially for the chemical composition and the vertical distribution) is too low to reduce the uncertainties apart from this small set. Roughly speaking, on a regular basis and at European scale, a few hundreds of measurements are available (ground stations for ozone, aerosols and a few other chemical species). This should be compared with 10^6 or 10^7 concentrations tracked in time by the model.

The use of multimodel approaches is therefore a promising tool. As said before, the modeling system POLYPHEMUS (http://cerea.enpc.fr/polyphemus/) has been designed for this aim. Each model used by POLYPHEMUS

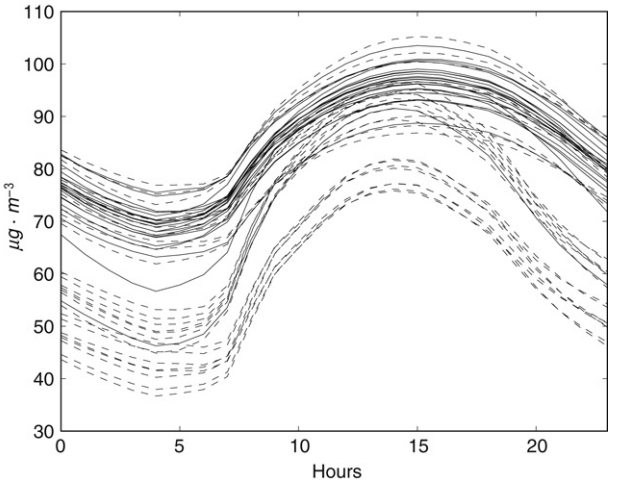

Fig. 3. Ozone daily profiles from 48 models built from combinations of physical parameterizations, data bases and numerical algorithms. Each hourly value of the profile is the average of ozone concentrations over the whole domain (Western Europe) and all days (summer 2001).

is defined as a configuration of parameterizations, a choice of data bases and a choice of numerical algorithms and discretization levels.

Let $\mathcal{E}$ be the set of resulting models, labelled by m. The objective of ensemble forecast is to compute a forecast $M_{t,x}$ on the basis of the model outputs $M_{m,t,x}$ for model m (at time t and at position x).

A classical approach is of course to use the ensemble mean (EM in the sequel):

$$\mathrm{EM}_{t,x} = \frac{1}{|\mathcal{E}|} \sum_{m \in \mathcal{E}} M_{m,t,x}. \tag{4}$$

We refer for instance to [14] for other "simple" approaches (the application is the forecast of radionuclides).

When observations are available, other approaches may be used. Let $O_{t,x}$ be the observation corresponding to the model output $M_{t,x}$. The forecast may be given as a linear combination of model outputs as:

$$\mathrm{ELS}_{t,x} = \sum_m \alpha_m M_{m,t,x} \tag{5}$$

where the vector α minimizes the cost function $\sum_{t,x} \left[O_{t,x} - \sum_m \alpha_m M_{m,t,x} \right]^2$ which measures the model-to-data discrepancy. The minimization can be performed in the least square sense (ELS stands for ensemble least square). This technique is sometimes referred as *superensemble* [15].

In order to improve the combination, one may use time-dependent weights:

$$\mathrm{ELS}^{\mathrm{d}}_{t,x} = \sum_m \alpha^{\mathrm{d}}_{m,t} M_{m,t,x} \tag{6}$$

where the upper script, d stands for "date" (time dependence). The optimal weights $\alpha^{\mathrm{d}}_{m,t}$ minimize:

$$\sum_x \left[O_{t,x} - \sum_m \alpha^{\mathrm{d}}_{m,t} M_{m,t,x} \right]^2. \tag{7}$$

We kept the same weights for all stations so that the weights can hopefully be applied everywhere in the domain. This way, one may still forecast 2D fields, which is one of the main features of an air quality model as compared to purely statistical models. Nevertheless, in the following paragraphs, the weights are predicted based on past observations from all observation locations and the forecast skills of the combinations are assessed at the same locations.

A key point is of course to forecast the weights α using only *past* observations. An example is the forecast of ozone peaks. Unfortunately, the optimal weights are very hard to forecast because of their erratic time evolution (Fig. 4).

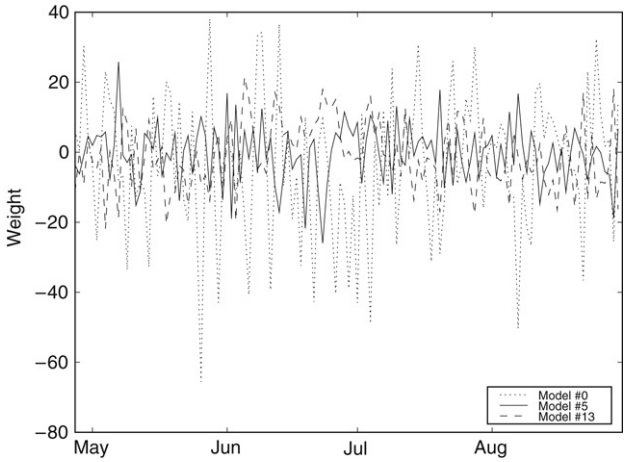

Fig. 4. Time evolution of the three weights of $\mathrm{ELS}^{\mathrm{d}}$.

Table 3
Performances (RMSE in $\mu g\ m^{-3}$) of the best model, $\mathrm{ELS}^{\mathrm{d},30}$, ELS and $\mathrm{ELS}^{\mathrm{d}}$

Combination	Hourly	Peak
Best model	28.5	23.9
$\mathrm{ELS}^{\mathrm{d},30}$	22.8	21.2
ELS	22.9	20.2
$\mathrm{ELS}^{\mathrm{d}}$	15.3	12.4

Performances of the ensemble mean EM are similar to that of the best model.

The weights may be computed over a moving learning period of 30 observational steps (that is, with the observations of ozone peaks from the 30 previous days):

$$\mathrm{ELS}^{\mathrm{d},30}_{t,x} = \sum_m \alpha^{30}_{m,t} M_{m,t,x} \tag{8}$$

where $\alpha^{30}_{m,t} = (\alpha^{30}_{1,t}, \alpha^{30}_{2,t}, \alpha^{30}_{3,t}, \ldots)$ minimizes:

$$\sum_{t'=t-30}^{t'=t-1} \sum_x \left[O_{t',x} - \sum_m \alpha^{30}_{m,x} M_{m,t',x} \right]^2. \tag{9}$$

The performances of this algorithm, applied to ozone peaks and to ozone hourly concentrations, are illustrated in Table 3. The combination $\mathrm{ELS}^{\mathrm{d},30}$ gives significant improvements as compared to the best model in the ensemble. Nevertheless there is still a lot of work to get the optimal weights ($\mathrm{ELS}^{\mathrm{d}}$).

For this purpose, there are promising algorithms from the machine learning field. Our tests show that the most simple algorithms from machine learning (such as the gradient descent algorithm — [16]) do not perform well: they lead to the same performances as the best model or the ensemble mean.

3. Uncertainties and data assimilation

A field of growing interest in air quality modeling is related to data assimilation and inverse modeling (e.g., [17]). When observational data provided by monitoring networks are available, the coupling between this data and the model outputs can reduce the model uncertainties either for forecast as such (data assimilation) or for some input parameters (inverse modeling of emissions for instance).

We use once more the input/output formulation $y = F(\Psi)$. These techniques are based on the minimization of the cost function

$$J_o(\Psi) = \sum_{i=1}^{i=n} (\text{obs}_i - H_i\, F_i(\Psi))^T R_i^{-1} (\text{obs}_i - H_i\, F_i(\Psi)) \tag{10}$$

with H_i the so-called observation operator at time t_i, obs_i an observation (n is the number of observations), F_i the corresponding model output and R_i the observation covariance matrix. A background term can be added to this cost function:

$$J_b(\Psi) = (\Psi - \Psi_b)^T B^{-1} (\Psi - \Psi_b). \tag{11}$$

This term takes into account a first estimate Ψ_b of Ψ (a *background* value). Mathematically speaking, this can be viewed as a penalty term (Tikhonov regularization).

The data assimilation/inverse modeling problem results in the search for:

$$\Psi^\star = \text{argmin}_{\Psi \in V_{\text{adm}}} J \stackrel{\Delta}{=} J_o + J_b \tag{12}$$

where V_{adm} is the space of admissible input parameters (for instance a set of positive parameters).

Let us assume that other uncertain input parameters are not included in Ψ (note that, due to the small set of observational data, the minimization problem is usually ill-conditioned: the dimension of Ψ should not be too large). We now write $y = F(\Psi, \Phi)$ with Φ the other parameters. The optimized inputs $\Psi^\star$ are a function of Φ:

$$\Psi^\star(\Phi) = \text{argmin}_\Psi J(\Psi, \Phi). \tag{13}$$

In a typical application, Ψ stands for the emissions and Φ for parameters used in physical parameterizations or meteorological fields.

A sensitivity analysis (similar to what we have already presented) can be performed. The usual terminology is "second-order sensitivity" [18]. By using the implicit function theorem, the differentiation of the optimal conditions $\nabla_\Psi J = 0$ that define $\Psi^\star$ leads to:

$$\frac{\partial \Psi}{\partial \Phi} = -\,(\text{Hess}\, J)^{-1}\, \frac{\partial \nabla_\Psi J}{\partial \Phi} \tag{14}$$

with Hess J the Hessian matrix of J.

We refer for instance to [19] with the application to a Gaussian model used for operational forecast of short-range dispersion of radionuclides. In [20], a brute-force approach is used to estimate the impact of uncertainties on optimized emissions of nitric oxide and nitrogen dioxide. A few parameterizations are modified and the resulting optimized parameters are compared.

These methods only indicate a local sensitivity. Global methods can also be used and the coupling to ensemble methods is a promising approach.

4. Conclusions and future works

The topics related to uncertainties are of growing interest in air quality modeling because models are highly uncertain while used in real life applications. Even if models are usually presented as validated thanks to model-to-data comparisons, the set of available observational data is too small to control the uncertainties. Appropriate techniques, such as Monte Carlo simulations or sensitivity analysis, show a wide spread in model outputs.

The use of a tuned model with respect to a given target is therefore not relevant for impact or screening studies (where the conditions in which the model was tuned are not met anymore). The use of multi-models approaches are necessary and promising. We have proposed in this paper a few techniques that can be used for ensemble forecasting based on observational data (superensemble with forecasted weights).

Many points are still unresolved. At a numerical level, the development of efficient Monte Carlo methods is still an issue because of the computational burden for high-dimensional systems arising in air quality modeling. The appropriate coupling between ensemble methods and data assimilation is another difficult issue which requires a methodological framework and the development of dedicated numerical algorithms.

References

[1] W.R. Stockwell, F. Kirchner, M. Kuhn, S. Seefeld, A new mechanism for regional atmospheric chemistry modeling, J. Geophys. Res. 102 (D22) (1997) 25847–25879.

[2] A. Pourchet, V. Mallet, D. Quélo, B. Sportisse, Some numerical issues in Chemistry-Transport Models — A comprehensive study with the Polyphemus/Polair3D platform, Tech. Rep. 26, CEREA, 2005.

[3] V. Mallet, Estimation de l'incertitude et prévision d'ensemble avec un modèle de chimie-transport – Application à la simulation numérique de la qualité de l'air, Ph.D. Thesis, École nationale des ponts et chaussées, 2005.

[4] S.R. Hanna, J.C. Chang, M.E. Fernau, Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM–IV) due to uncertainties in input variables, Atmos. Environ. 32 (21) (1998) 3619–3628.

[5] S.R. Hanna, Z. Lu, H.C. Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, 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. Environ. 35 (5) (2001) 891–903.

[6] M.A. Tatang, Direct incorporation of uncertainty in chemical and environmental engineering systems, Ph.D. Thesis, Massachusetts Institute of Technology, 1995.

[7] J. Boutahar, Méthodes de réduction et de propagation d'incertitudes: Application à un modéle de chimie-transport pour la modélisation et la simulation des impacts, Ph.D. Thesis, École nationale des ponts et chaussées, 2004.

[8] V. Mallet, B. Sportisse, 3-D chemistry-transport model Polair: Numerical issues, validation and automatic-differentiation strategy, Atmos. Chem. Phys. Discuss. 4 (2004) 1371–1392.

[9] R. Djouad, N. Audiffren, B. Sportisse, A sensitivity analysis study for RADM2 mechanism using automatic differentiation, Atmos. Environ. 37 (22) (2003) 3029–3038.

[10] V. Mallet, B. Sportisse, A comprehensive study of ozone sensitivity with respect to emissions over Europe with a chemistry-transport model, J. Geophys. Res. 110 (D22) (2005).

[11] J.A. Hoeting, D. Madigan, A.E. Raftery, C.T. Volinsky, Bayesian model averaging: A tutorial, Stat. Sci. 14 (4) (1999) 382–417.

[12] V. Mallet, D. Quélo, B. Sportisse, M. Ahmed de Biasi, É. Debry, I. Korsakissok, L. Wu, Y. Roustan, K. Sartelet, M. Tombette, H. Foudhil, Technical Note: The air quality modeling system Polyphemus, Atmos. Chem. Phys. 7 (20) (2007).

[13] V. Mallet, B. Sportisse, Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling, J. Geophys. Res. 111 (D1) (2006).

[14] S. Galmarini, R. Bianconi, W. Klug, T. Mikkelsen, R. Addis, S. Andronopoulos, P. Astrup, A. Baklanov, J. Bartniki, J.C. Bartzis, et al., Ensemble dispersion forecasting —- part I: Concept, approach and indicators, Atmos. Environ. 38 (28) (2004) 4607–4617.

[15] T.N. Krishnamurti, C.M. Kishtawal, Z. Zhang, D.B.T. LaRow, E. Williford, Multimodel ensemble forecasts for weather and seasonal climate, J. Clim. 13 (2000) 4196–4216.

[16] N. Cesa-Bianchi, P.M. Long, M.K. Warmuth, Worst-case quadratic loss bounds for prediction using linear functions and gradient descent, IEEE Trans. Neural Net. 7 (3) (1996) 604–619.

[17] H. Elbern, H. Schmidt, Ozone episode analysis by four-dimensional variational chemistry data assimilation, J. Geophys. Res. 106 (D4) (2001) 3569–3590.

[18] Z. Wang, I. Navon, F.-X. Le Dimet, X. Zou, The second order adjoint analysis: Theory and applications, Meteorol. Atmos. Phys. 50 (1992) 3–20.

[19] D. Quélo, B. Sportisse, O. Isnard, Data assimilation for short range atmospheric dispersion of radionuclides: A case study of second-order sensitivity, J. Environ. Radioact. 84 (3) (2005) 393–408.

[20] D. Quélo, V. Mallet, B. Sportisse, Inverse modeling of NO_x emissions at regional scale over northern France: Preliminary investigation of the second-order sensitivity, J. Geophys. Res. 110 (D24) (2005).