RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **96** (3), 2002, pp. 507–528
Matemática Aplicada / Applied Mathematics
Artículo panorámico / Survey

Some tracks in Air Pollution Modelling and Simulation

B. Sportisse, J. Boutahar, E. Debry, D. Quélo and K. Sartelet

Abstract. In this article we discuss some issues related to Air Pollution modelling (as viewed by the authors): subgrid parameterization, multiphase modelling, reduction of high dimensional models and data assimilation. Numerical applications are given with POLAIR, a 3D numerical platform devoted to modelling of atmospheric trace species.

Algunos resultados para el modelado y simulación de la polución aérea

Resumen. En este artículo se presentan algunos temas relacionados con polución aérea, reducción de modelos de dimensiones grandes y asimilación de datos. Se dan aplicaciones numéricas con POLAIR: una plataforma numérica dedicada al modelado de trazas atmosféricas de especies.

1. Introduction

Description of the chemical state of the atmosphere is now a key issue for many environmental applications, ranging from global scale (greenhouse effect) to regional (transboundary pollution) or local scales (photochemical smog).

modelling is a powerful tool in addition to field campaigns and laboratory experiments. Many Chemistry-Transport-Models (CTMs) are now available. The underlying models are given by large sets of coupled Partial Differential Equations that describe the time and space evolution of trace species in the phases to be considered ([49] for an overview). If c_i stands for the concentration of chemical gas-phase species labelled by i:

$$\frac{\partial c_i}{\partial t} + \underbrace{div(\vec{V}(\vec{x},t)c_i)}_{\text{wind advection}} = \underbrace{div(K(\vec{x},t)\nabla c_i)}_{\text{turbulent diffusion}} + \underbrace{\chi_i(\vec{c},T(\vec{x},t),RH(\vec{x},t),I(\vec{x},t))}_{\text{chemical production}} + \underbrace{S_i(\vec{x},t)}_{\text{source}} + Param \tag{1}$$

where $\vec{V}$ (the wind velocity), K (the eddy diffusivity matrix), T (the temperature), I (the actinic flux describing the radiative state of the atmosphere) and RH (the relative humidity) are fields given by meteorological models. In the above equation, χ_i stands for *local* chemical transformations. $Param$ stands for the parameterizations (see below, e.g. scavenging by rain drops).

Presentado por Haïm Brezis.
Recibido: 26 de Julio de 2002. *Aceptado*: 9 de Octubre de 2002.
Palabras clave / Keywords: Numerical simulation, Air Pollution modelling
Mathematics Subject Classifications: 65M06, 65M20, 65Y05, 65Y20

Notice that there is usually no feedback from the chemical state to the meteorological state, even if this should be taken into account, especially trough radiative transfer (I influences the photolytic reactions).

Boundary conditions have of course to be added to Eq. (1). A key boundary condition is at ground:

$$-K_z \frac{\partial c_i}{\partial z} = \underbrace{E_i(\vec{x},t)}_{\text{emission}} - \underbrace{v_i^{dep}(\vec{x},t)c_i}_{\text{dry deposition}} \tag{2}$$

where K_z stands for the vertical eddy diffusivity (given by a parameterization, for instance [30]), E_i is an emission factor and v_i^{dep} is a dry deposition velocity ([70]).

Gas-phase chemical mechanisms are now more or less well understood and take into account up to hundreds of species ([18]). In such models, the term χ_i is directly related to chemical kinetics. One of the key challenge of current atmospheric modelling is the fine description of multiphase mixture by taking into account aqueous-phase processes (for instance inside clouds) or particles (aerosols). The trace species are then not only in gaseous phase but also in aqueous phase or in aerosols.

The condensed matter indeed has various influences on the atmospheric state through mass transfer and radiative effects. modelling of aerosols is a challenging task due to the need for size-resolved models, giving not only the chemical composition of dissolved species but also the distribution of particles with respect to size (volume or radius).

If $n(v,t)$ is the volume distribution of a (let say a *single* component) aerosol, the General Dynamic Equation (GDE) reads ([49]):

$$\frac{\partial n}{\partial t} = \underbrace{\frac{1}{2}\int_0^v K(v-q,q)n(v-q,t)n(q,t)dq - n(v,t)\int_0^\infty K(q,v)n(q,t)dq}_{\text{coagulation}}$$

$$\underbrace{-\frac{\partial}{\partial v}(I(v)n(v,t))}_{\text{growth}} + \underbrace{J_0(v)\delta_(v_0)(v)}_{\text{nucleation}} \tag{3}$$

Many processes have to be taken into account:

- coagulation between particles, described by the coagulation kernel K,
- growth by condensation/evaporation of some gas-phase species onto existing particles (gas to particle conversion) described by the growth rate of volatile species I, given by thermodynamics (see below),
- nucleation of the smallest aerosols from stable aggregates of clusters: v_0 is the nucleation threshold (typically 40 nanometers) and J_0 the nucleation rate.

Apart from a lack of physical knowledge (especially for nucleation, [22]), this model illustrates why aerosol modelling is a computationally intensive and difficult task because of the number of processes involved.

In practice *multi-component* models are used. Aerosols are either *dry* or *wet* particles and are composed by three types of chemical constituents: solids, dissolved species and ions. The chemical speciation inside aerosols has to be computed and one defines $q_i(v)$ the mass distribution of a "species" i inside an aerosol of size v. Equations similar to the GDE may be given for the multi-component case, which drastically increase the complexity.

For instance, the growth rate I_i for species i (e.g. $(NH_3)_{aq}$) is defined by thermodynamics through:

$$I_i(v) = \lambda_i(v)(c_i - c_i^{eq}(q_1,\ldots,q_s)) \tag{4}$$

where c_i is the concentration of the related gas-phase species (e.g. $(NH_3)_g$) and λ_i a mass transfer coefficient. $c_i^{eq}(m_1,\ldots,m_s)$ is defined as the value at thermodynamical equilibrium of the aerosol internal

chemistry (with s species, typically 20) and is computed by a thermodynamical module ([38, 71]) solving large differential-algebraic equations (in each grid cell, for each size of the distribution!). CPU costs can therefore drastically increase and most of the CPU time needed for aerosol models is related to thermodynamics.

This brief overview of the underlying models to be handled with in atmospheric chemistry illustrates the increase of the computational burden in current CTMs. The key issues are dominated by the following considerations: these models are highly non linear, coupled, of high dimension and with a large range of spatial and time scales.

This article investigates these issues. Subgrid parameterizations are presented in the first section, with a particular emphasis on mass transfer (scavenging) between gas-phase species and liquid phase (cloud drops or rain).

Numerical issues are investigated in the second section with "classical" topics: solvers for "stiff" multiphase models and splitting methods. A key point is the need for appropriate *reducing* techniques.

Many input parameters are known with a poor accuracy. Data assimilation and inverse modelling are therefore challenging topics of atmospheric chemistry; many numerical difficulties therefore occur, especially with the use of the so-called adjoint models. These points are described in the third section.

We end this reviewing paper with the example of a state-of-the-art 3D Eulerian Chemistry-Transport-Model, namely POLAIR, and some applications.

2. Some examples of parameterizations

2.1. Some models

The horizontal dimension of a grid cell in a current CTM ranges typically from 1 kilometer (ozone forecast at the urban scale) to 100 kilometers (regional transport of transboundary pollutants, acid rains).

Because many physical processes occur at scales smaller than 1 kilometer, subgrid parameterizations is necessary in CTMs. We can cite for instance the following examples:

- *Convective processes*:

  Inside regions of strong gradients of humidity, updrafts and downdrafts may occur. This implies a modification of the vertical distribution of pollutants. Many parameterizations ([63]) have been proposed and are more or less related to the following description of convective process:

  $$\frac{dc(z_j)}{dt} = \sum_{k=1}^{k=n_z} T_{jk}^{conv} c(z_k) \tag{5}$$

  with (z_k) the vertical discretized grid, T_{jk}^{conv} an exchange term between levels j and k, parameterizing downdrafts and updratfs. Notice that convective "diffusion" takes into account (*a priori*) the whole column on the contrary to the "classical" (3 points) diffusion. A rigorous derivation of such parameterizations is still an open question.

- *Plume-in-grid models*:

  Point sources (S_i in Eq. (1)) are typically related to plants. Generating a diffuse source over a grid cell of dimension 100 kilometers is of course far from reality. Plume-in-grid models (PIG) use a parameterization of the dispersion of the plume inside the cell in order to circumvent this difficulty; the key point is to derive a simple (Gaussian-like) model on the basis of the available meteorological data in the cell (one value for each field).

  We refer for instance to [23] for an overview of current techniques.

  Notice that from a numerical point of view, an alternative technique is the use of adaptive gridding techniques in order to give a finer description of the flow in the vicinity of the sources ([64, 58]).

- *Segregation effects:*

  In the same vein, the interaction between chemistry and turbulence is a key point, which is far from being addressed in current CTMs. If $c(x,t)$ is the local (true) concentration and $< c(t) >$ the average concentration in a given grid cell Ω, fluctuations are defined by $c^{'}(x,t) =< c(t) > -c(x,t)$.

  The chemical mechanism is defined for the "true" concentrations and for bimolecular reactions one has typically:

$$\frac{dc}{dt} = -kc^2 \tag{6}$$

  with k a kinetic rate. After averaging over the grid cell:

$$\frac{d < c >}{dt} = -\frac{k}{Vol(\Omega)} \int_{\Omega} c(x,t)^2 dx = -k(< c >^2 + < (c^{'})^2 >) \tag{7}$$

  In CTMs, an assumption of a well-mixed reactor is done and the chemical mechanism is used as:

$$\frac{d < c >}{dt} = -k < c >^2 \tag{8}$$

  and the error is directly related to the so-called *segregation intensity*:

$$I_s = \frac{< (c^{'})^2 >}{< c >^2} \leq 1 \tag{9}$$

  For large values of I_s, this model overestimates the chemical production rate. We refer to [52] for the application to the case of box models and to [36] for the a particular case of chemistry/dynamics interactions.

- *Mass transfer and multiphase processes*:

  The typical radius of a cloud droplet is 10 micrometers. During cloudy events, mass transfer between the gas-phase and the liquid phase may occur for soluble species. Inside droplets, aqueous-phase chemistry has to be taken into account and may influence the gas-phase species through mass transfer. Such (nonlinear) phenomena occurring at these small scales have to be modelized in CTMs. These processes are usually referred as *in-cloud scavenging*.

  During rainy events, gas-phase species may be also scavenged by falling raindrops. This is the *below-cloud scavenging*. The typical size is of magnitude 1 millimeter and the same parameterization issue has to be addressed.

  In the same vein, the mass transfer between the gas-phase and aerosols (condensation/evaporation) occurs at scales of magnitude ranging from 100 nanometers to $1-5$ micrometers (typical spectrum for atmospheric particles). These processes have a key impact on the gas-phase through the formation of the so-called Secondary Organic Aerosols (SOA, [39]).

The next section focuses on wet scavenging.

2.2. Parameterization of wet scavenging

2.2.1. Below-cloud scavenging ([55])

We consider a sample model of a mono-disperse rain: falling raindrops are supposed to have a fixed radius a and a falling (vertical) velocity $U(a)$ (in m/s), given by an appropriate parameterization as a function of a.

Let c_g (resp. c_a) be the concentration of a gas-phase soluble species (resp. the associated liquid-phase species). c_a is defined with respect to the liquid phase. The mass transfer between both phases is then given

by an advection-reaction model:

$$\begin{cases} \dfrac{\partial c_g}{\partial t} = -\dfrac{L}{\tau_{transfer}} \left(c_g - \dfrac{c_{aq}}{HRT} \right) \\ \dfrac{\partial c_{aq}}{\partial t} + U \dfrac{\partial c_{aq}}{\partial z} = \dfrac{1}{\tau_{transfer}} \left(c_g - \dfrac{c_{aq}}{HRT} \right) \end{cases} \tag{10}$$

Notice that chemistry is neglected. z is the fall distance (below the cloud at height h, with the convention $U > 0$). H is the Henry's constant defined by thermodynamical equilibria of mass transfer (to be given for each species), R is the ideal gas constant and T the temperature. L is the liquid water volume fraction (the ratio of liquid water volume to air volume) defined by:

$$L = \frac{10^{-6}}{3.6} \frac{p_0}{U} \tag{11}$$

with p_0 the rain intensity (in mm/h). $c_{aq} = L c_a$ is defined with respect to the gas phase. $\tau_{transfer}$ is the characteristic time for mass transfer and is defined by some micro-physical parameters.

It is of course not possible to solve such models in the framework of CTMs and parameterizations are used. A classical parameterization is to ignore aqueous-phase concentrations (inside rain droplets) and to add a new source term for gas-phase species under the form:

$$\frac{dc_g}{dt} = -\Lambda_g c_g \tag{12}$$

where $\Lambda_g(p_0, a, HRT, \tau_{transfer})$ is the so-called scavenging coefficient.

On the basis of asymptotic expansions ([65]), [55] discusses the range of validity of such parameterizations. The key parameters are:

$$\epsilon_1 = L \frac{\tau_{fall}}{\tau_{transfer}}, \quad \epsilon_2 = L\, HRT \tag{13}$$

with $\tau_{fall} = h/U$ the fall time for a droplet. ϵ_1 measures the ability of a falling drop to exchange mass with gas-phase during its fall. ϵ_2 indicates the validity of equilibrium between gaseous and aqueous phases.

The parameterization is good for species easy to be dissolved such as HNO_3 while it may be less accurate for other species like SO_2 if the rain intensity is large and for small raindrops.

Many open points need still to be addressed, especially the extension to polydisperse raindrops.

2.2.2. In-Cloud scavenging ([54])

In-Cloud scavenging is a similar topic. The impact of clouds on gas-phase species recover different phenomena:

- a dynamical interaction through convective processes (as indicated above, see for instance [63]),
- radiative forcing: cloud properties may deeply modify the actinic flux available for photolytic reactions below the cloud ([31]),
- mass transfer inside the cloud between the droplets and the gas-phase (Figure 1).

This latter phenomena may have a strong impact ([29]) and has to be parameterized in CTMs.

We consider a fixed cloud droplet Ω_a surrounded by gas-phase, supposed to be a sphere (Figure 2). The gas-volume, Ω_g (supposed to be a sphere as well), is such that:

$$\frac{Vol(\Omega_a)}{Vol(\Omega_g)} = L \tag{14}$$

where L is the liquid water content (v/v). This is the so-called *unique droplet model.*

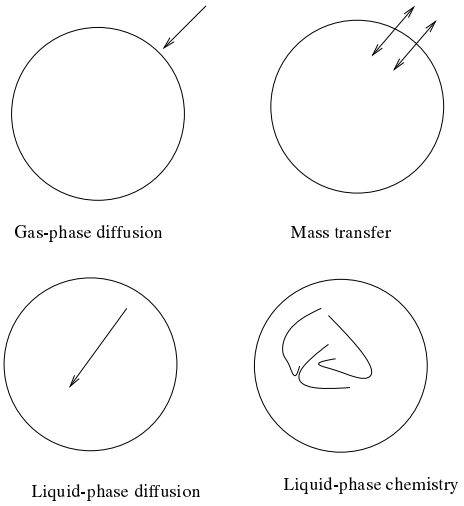

Figure 1. In-Cloud scavenging

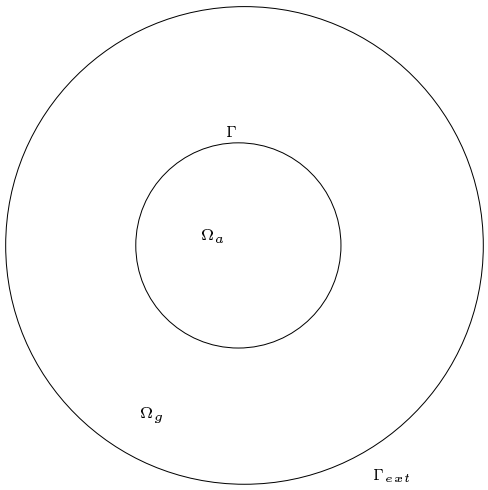

Figure 2. Unique droplet model

We use the same notations as before. The mass transfer problem is:

- for the gas-phase (over Ω_g):

$$\frac{\partial c_g}{\partial t} = D_g \Delta c_g + \chi_g(c_g) \tag{15}$$

- for the dissolved species (inside the droplet):

$$\frac{\partial c_a}{\partial t} = D_a \Delta c_a + \chi_a(c_a) \tag{16}$$

where D_a and D_g are respectively the molecular diffusions for aqueous and gaseous species. χ_g (resp. χ_a) stands here for the chemical production in gas-phase (resp. in aqueous phase). Let us precise the boundary conditions:

- There is a no-flux boundary condition on Γ_{ext}, the outer boundary layer for the gas-phase:

$$D_g \frac{\partial c_g}{\partial n} = 0 \quad \text{on } \Gamma_{ext} \tag{17}$$

- At the droplet surface, the fluxes are equal and describe a relaxation to Henry's equilibrium:

$$D_a \frac{\partial c_a}{\partial n} = D_g \frac{\partial c_g}{\partial n} = D_a \frac{\partial c_a}{\partial n} = \frac{1}{4} \alpha \, v (c_g - \frac{c_a}{HRT}) \tag{18}$$

  where α is the *accommodation coefficient*, v is the thermal velocity.

Remember that because the typical size for a cloud droplet is 10 micrometers, a parameterization is necessary.

We define the *bulk* gas-phase and liquid-phase concentrations by:

$$\bar{c_a} = \frac{1}{\Omega_a} \int_{\Omega_a} c_a(x) dx \ , \ \ \bar{c_g} = \frac{1}{\Omega_g} \int_{\Omega_g} c_g(x) dx \tag{19}$$

We prove in [54] the validity of the following parameterization:

$$\left\{ \begin{array}{l} \frac{d\bar{c}_g}{dt} = \chi_g(c_g) - L \, k_{mt}(a)(c_g - \frac{c_a}{HRT}) \\ \frac{d\bar{c}_a}{dt} = \chi_a(c_a) + k_{mt}(a)(c_g - \frac{c_a}{HRT}) \end{array} \right. \tag{20}$$

where the mass transfer coefficient k_{mt} is defined in the following way:

$$k_{mt}(a) = \frac{1}{\frac{a^2}{3D_g} + \frac{4a}{3v\alpha}} \tag{21}$$

for large diffusivities (which is indeed met in practice). We write τ_{dg} (resp. τ_{da}) the time scale associated with gas-phase diffusion (resp. aqueous phase diffusion), namely:

$$\tau_{dg} = \frac{a^2}{D_g} \ , \ \ \tau_{da} = \frac{a^2}{D_a} \tag{22}$$

The key point is that the homogenization timescales τ_{dg} and τ_{da} have small values. If τ_{χ_g} (resp. τ_{χ_a}) is a characteristic timescale for χ_g (resp. χ_a) and:

$$\epsilon_{dg} = \frac{\tau_{dg}}{\tau_{\chi_g}} \ , \ \ \epsilon_{da} = \frac{\tau_{da}}{\tau_{\chi_a}} \ , \tag{23}$$

the parameterized model is valid for $\epsilon_{dg} \sim 0$ and $\epsilon_{da} \sim 0$.

We refer to [7, 17, 37] for some related mathematical problems (and the so-called "lumped parameter assumption"). A physical interpretation may be found in [48].

An open question is the extension of the above analysis for the case of fast chemistry. As for below-cloud scavenging, the extension to polydisperse clouds has to be done. Surface chemistry is also a topic of growing interest with highly sophisticated models.

3. Numerical modelling of multiphase atmospheric flows

Suppose that a model relying on Eq. (1) has been chosen on the basis of appropriate parameterizations as given in Section 2..

Many numerical issues have to be tackled with in order to solve such equations, especially in a forecast framework, for which CPU time has to be drastically reduced. We refer to [66] for an overview of the numerical simulation of Air Pollution.

We will try in this section to highlight some typical "difficult" points:

- the use of splitting methods,
- the use of stiff solvers (many timescales have to be dealt with),
- the way size-resolved multiphase models may be solved.

To conclude, an alternative approach is to build *reduced models*.

3.1. Splitting methods

The model given by Eq. (1) takes into account many processes. As implicit solvers are used, numerical coupling between all the processes will lead to inversion of systems of large dimension (typically the product of the number of grid cells by the number of chemical species).

Splitting is then usually advocated ([34, 66]) in order to uncouple local reactive phenomena (one chemical box model per grid cell) and passive transport (one advection-diffusion problem per chemical species). Another pragmatic point of view is the possibility to use sub-models as black boxes. In practice, advection, diffusion and chemistry are solved in a splitting sequence.

For the simple linear case with two processes A and B, consider the following evolution equation from time t_n (here c^n stands for the approximate numerical value of c at time t_n):

$$\frac{dc}{dt} = Ac + Bc\ ,\ \ c(0) = c^n \tag{24}$$

The first-order splitting method $(A-B)$ is defined by the following algorithm:

1. Integrate operator A: $\frac{dc^*}{dt} = Ac^* \quad \text{on} \quad [0, \Delta t]\ ,\ \ c^*(0) = c^n$,
2. Integrate operator B: $\frac{dc^{**}}{dt} = Bc^{**} \quad \text{on} \quad [0, \Delta t]\ ,\ \ c^{**}(0) = c^*(\Delta t)$
3. Output: $c^{n+1} = c^{**}(\Delta t)$.

Numerics associated to operator splitting is well known (see [32] for instance). The error is related to the commutator $AB - BA$ in the previous case and we refer to [20, 27, 47] for an extension to the nonlinear case and the use of the Baker-Campbell-Hausdorff formula.

In the specific case of Air Pollution modelling, many other algorithms have already been used. For (second-order) Strang splitting ([60]) the sequence reads:

1. Integrate B: $\frac{dc^*}{dt} = Bc^* \quad \text{on} \quad [0, \frac{\Delta t}{2}]\ ,\ \ c^*(0) = c^n$
2. Integrate A: $\frac{dc^{**}}{dt} = Ac^{**} \quad \text{on} \quad [0, \Delta t]\ ,\ \ c^{**}(0) = c^*(\frac{\Delta t}{2})$
3. Integrate B: $\frac{dc^{***}}{dt} = Bc^{***} \quad \text{on} \quad [0, \frac{\Delta t}{2}]\ ,\ \ c^{***}(0) = c^{**}(\Delta t)$
4. Output: $c^{n+1} = c^{***}(\frac{\Delta t}{2})$.

Source-splitting methods are an alternative approach. The idea is to avoid artificial transient layers by adding explicit source terms instead of changing initial conditions:

1. Integrate A: $\frac{dc^*}{dt} = Ac^* \quad$ on $\quad [0, \Delta t]\ , \quad c^*(0) = c^n$
2. Integrate operator B with a complementary source term:

$$\frac{dc^{**}}{dt} = Bc^{**} + \frac{c^*(\Delta t) - c^n}{\Delta t} \quad \text{on} \quad [0, \Delta t]\ , \quad c^{**}(0) = c^n \tag{25}$$

3. Output: $c^{n+1} = c^{**}(\Delta t)$.

Note that the initial condition for the second step has not been modified, such as avoiding any transient layer if B is a stiff model. A classical analysis proves that this is a first-order model.

Approximate Matrix Factorizations perform the splitting at the algebraic level. If an implicit algorithm is used for Equation (24) inversions of matrices similar to $I - (A + B)\Delta t$ has to be done. AMF methods are based on the following approximation:

$$I - (A + B)\Delta t \sim (I - A\Delta t)(I - B\Delta t) + O(\Delta t^2) \tag{26}$$

which avoids to invert simultaneously A and B. We refer for instance to [66, 4] for some illustrations.

We want here to focus on a particular point, namely the influence of the sequence order when stiff models have to be solved (see next section). We refer to [51, 68] for a deeper understanding.

The key idea is that the condition $\Delta t \to 0$ is not met in practice for stiff models as $\Delta t \gg \tau_{fast}$ where τ_{fast} corresponds to the fastest phenomena (that is exactly why implicit methods are used !). The classical analysis in order to derive convergence orders for the algorithm is then no more valid and an *order reduction* ([10]) of many of the above algorithms has to be taken into account.

Consider for instance a model with A related to slow phenomena (typically advection and diffusion in our case) and B related to slow AND fast phenomena (typically chemical kinetics). A key result (see subsection 3.4.) is that B is related to a low-dimensional model defined by equilibria of fast processes: the integration of B is associated with a projection to this low-dimensional model.

As as consequence, any sequence ending with the slow/fast operator B (in our case: chemistry) will conserve this property, which leads to a better accuracy of the splitting algorithm.

3.2. Stiff solvers

Atmospheric chemical kinetics is characterized by the wide range of characteristic timescales (over many decades): species have lifetimes that range from milliseconds and shorter (radicals, OH) to years (CH_4). This induces the well known stiffness of the resulting evolution equations and implicit solvers have to be used in order to perform time integration.

We refer for instance to [46] for a general benchmark of implicit and explicit solvers used in this field. For a given accuracy (of magnitude, let say, 1%) the second-order Rosenbrock method appears to outperform the other algorithms. For the integration of:

$$\frac{dc}{dt} = f(t, c) \tag{27}$$

the autonomous second-order Rosenbrock method is defined by the following steps ([67]) from time t_n to t_{n+1}:

$$c^{n+1} = c^n + \sum_{i=1}^{2} b_i\, k_i \tag{28}$$

where

$$\begin{cases} k_i = \Delta t\, f(t_n + \alpha_i\, \Delta t, c^n + \sum_{j=1}^{i-1} \alpha_{ij}\, k_j) + \Delta t\, J \sum_{j=1}^{i} \gamma_{ij}\, k_j \\ \alpha_i = \sum_{j=1}^{i-1} \alpha_{ij} \quad , \quad \gamma_i = \sum_{j=1}^{i} \gamma_{ij} \end{cases} \tag{29}$$

J denotes the Jacobian matrix $J = \partial f / \partial c$ and the formula coefficients b_i, α_{ij} and γ_{ij} are chosen to obtain a desired order of consistency and stability for stiff problems:

$$\left\{ \begin{array}{l} \gamma = 1 + \dfrac{1}{\sqrt{2}} \ , \ \gamma_1 = \gamma_{11} = \gamma \ , \ \gamma_{21} = -2\gamma \ , \ \gamma_{22} = \gamma \\ \gamma_2 = \gamma_{21} + \gamma_{22} = -\gamma \ , \ \alpha_1 = 0 \ , \ \alpha_2 = \alpha_{21} = 1 \ , \ b_1 = b_2 = \dfrac{1}{2} \end{array} \right. \tag{30}$$

We refer to [13] for the extension to the non-autonomous case with the particular application to cloudy events (strong gradients of the liquid water content have to be integrated).

3.3. Size-resolved models

As illustrated in the introduction the integration of the General Dynamic Equation (3) is a computationally difficult task.

A first family of algorithms is based on a stochastic formulation of (3) leading to Monte Carlo simulations: numerical particles represent the physical particles. This leads to expensive but accurate algorithms that may be used as reference solutions ([8]).

A second family of models is the so-called *modal approach*. Measured aerosol distribution appear to be a sum of independent families centered around a peak (Figure 3). A modal solution is then of the form:

$$n(D,t) = \sum_{i=1}^{p} n^{(i)}(\ln D) \ , \ \ n^{(i)}(\ln D) = \frac{N_i}{\sqrt{2\pi} \ln \sigma_i} \exp(-\frac{1}{2}(\frac{\ln D - \ln D_i}{\ln \sigma_i})^2) \tag{31}$$

The parameters (N_i, σ_i, D_i) uniquely define each mode and are functions of time. We refer to [5, 72, 1]

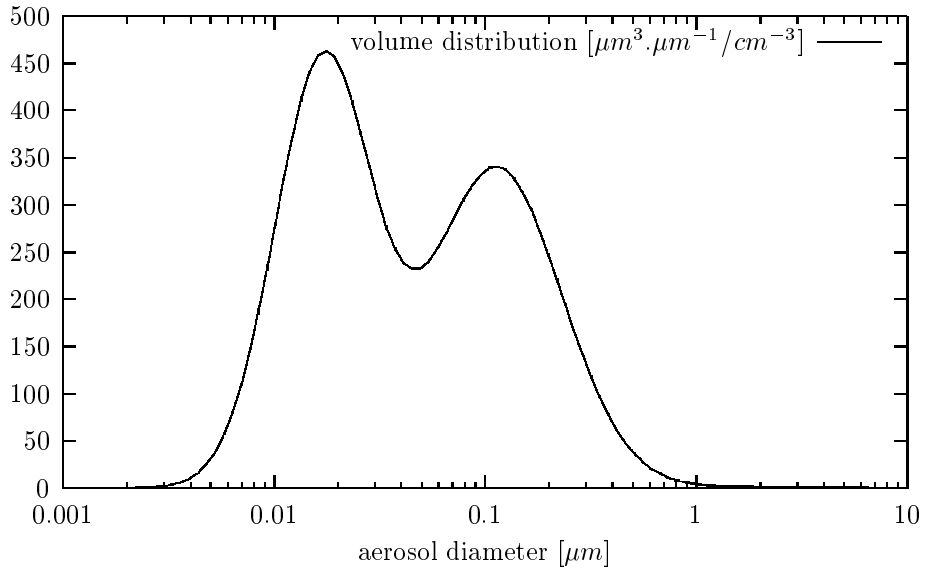

Figure 3. Modal distribution for continental aerosols (logarithmic scales)

for the definition of the appropriate evolution equation: after integration of the GDE under the above modal assumption, a closure model is found and thus leads to a loss of accuracy.

Size-resolved models do not make any *a priori* assumption for the aerosol distribution. Among them we first distinguish the sectional methods, which mainly consist in dividing the aerosol size spectrum into a given number of *bins*. These methods are widely used because of their roughness, especially to solve coagulation; we refer here to the so-called *size-binning* methods [21, 15]. The main drawbacks of such methods are still the lack of convergence results and the difficulty to extend it to the whole GDE.

Finite Element techniques (following [41]) may be also used to solve such models. For a multi-component model, one writes for species i:

$$v \in [v_0, \infty[\ ,\ \ q_i(v,t) = \sum_{j=1}^{p} q_i^j(t) L_j(v)\ ,\ \ q(v,t) = \sum_{j=1}^{p} q^j(t) L_j(v) \tag{32}$$

with $(L_j(v))$ a family of p given functions, e.g. Lagrange functions. With a Galerkin method, the set of unknown coefficients $Q = (q_i^j)_{(i,j)}(t)$ is the solution of an evolution equation:

$$\frac{dQ}{dt} = f(Q) \tag{33}$$

where f is related to the various processes of the GDE. We refer to [45, 9] for a deeper understanding of such models. Use of such models in current CTMs still remains an open issue.

3.4. Reducing methods

An alternative to the use of state-of-the art solvers is given by *reducing techniques*.

A first approach is related to the slow/fast behavior of chemical kinetics. The reduced model can be obtained by filtering the fast phenomena and the fast chemical species [53]. The second approach is related to the statistical behavior of these systems and is based on a proper orthogonal decomposition of the solutions [12]. The third method, High Dimensional Model representation, (HDMR,[44]), is an efficient representation of the input/output behavior of chemical models through look-up tables. It describes the output model by an expansion of finite hierarchical correlated function in terms of the input variables.

The physical model is written as $dc/dt = f(c)$ with $c \in R^n$ the vector of chemical concentrations.

3.4.1. Slow/fast reduction

The dynamical behavior of stiff systems is characterized by the existence of a fast transient phase followed by a slow long-term evolution. All the trajectories in the phase space of chemical concentrations (parameterized by time t) converge to an attracting manifold after a transient phase.

In practice we perform a splitting of species c into two sets: a set of *slow species* (c_s) and a set of *fast species* (c_f). The manifold is then defined by algebraic relations of the form $g(c_s, c_f) = 0$, which replace the evolution equations for fast species. The key point is that $g(c_s, c_f) = 0$ has to provide a way for computing c_f as a function of c_s ($c_f = h(c_s)$). Notice that this provides the theoretical background for Quasi-steady State Assumption ([19]).

A general technique is based on the linearization of the system: the search for the partitioning is then related to the search for the eigenvectors of the Jacobian matrix ([26, 33]). This may be related to the classical lumpings of species. We refer to [53] for a deeper understanding.

3.4.2. Proper Orthogonal Decomposition

The key idea of this method is to search for a basis which contains all information about the behavior of the exact solution. The reduced model is then the projection of the initial model in this basis.

In practice, we use the so-called "method of snapshots". It consists in computing an exact trajectory $(c(t))$. If there are N time steps (corresponding to times $t_1, \ldots, t_N$), we compute $\mathcal{C} = span(c(t_1), \ldots, c(t_N))$.

$\mathcal{C}$ may be viewed as a set of experimental data in R^n and the objective is to extract as many informations as needed from this set.

Of course all these informations are contained in a basis of $\mathcal{C}$, $(\Psi = (\Psi_1\ ,\ \Psi_2 \dots\ \Psi_d))$, whose dimension is d.

The POD methods is based on a particular choice of the orthonormal basis Ψ such that $c(t_j)$ can be *optimally* approached. Ψ is then such that for all $1 \leq k \leq d$, $(\Psi_1 \dots \Psi_k)$ minimizes:

$$J_k(\Phi) = \sum_{j=1}^{n} \|c(t_j) - \sum_{i=1}^{i=k} < \Phi_i, c(t_j) >_A \Phi_i\|_A^2 \tag{34}$$

with $< .,. >_A$ a particular choice of a scalar product. In practice $(\Psi_1 \dots \Psi_k)$ can be computed from the correlation matrix of the snapshots K ([25]).

The choice of a truncated basis of dimension p is related to the decreasing rate of the eigenvalues of the correlation matrix. p is then the number of degrees of freedom kept.

We now search for a solution $z(t) = \sum_{i=1}^{p} z_i(t)\ \Psi_i\ ,\ \ z_i \in R$ such that for $1 \leq i \leq p$:

$$\frac{dz_i}{dt} =< f(\sum_{i=1}^{p} z_i(t)\ \Psi_i), \Psi_i >_A\ \ ,\ \ z_i(0) =< c(0), \Psi_i >_A \tag{35}$$

3.4.3. HDMR Technique

Consider one call of the chemical module from time t_n to time t_{n+1} as a model with m input variables $x = (x_1, x_2 ..., x_m)$ and an output $c = F(x)$. x includes the initial concentrations and the appropriate forcing parameters used in chemistry (for example temperature).

HDMR expresses $F(x)$ as an expansion through:

$$\begin{aligned} F(x) =& f_0 + \sum_{i=1}^{i=m} f_i(x_i) + \sum_{1 \leq i < j \leq m} f_{ij}(x_i, x_j) \\ &+ \sum_{1 \leq i < j < k \leq m} f_{ijk}(x_i, x_j, x_k) + \dots + f_{1,2,\dots,m}(x_1, x_2, ..., x_n) \end{aligned} \tag{36}$$

where f_0 denotes a mean effect, $f_i(x_i)$, the effect of the input variable x_i acting independently and $f_{ij}(x_i, x_j)$ the cooperative effect of the input variables x_i and x_j. The last term $f_{1,2,\dots,m}(x_1, x_2, \dots, x_m)$ represents any residual dependence of all input variables acting together.

In practice, second-order expansions may provide a satisfactory description of $F(x)$ for many high dimensional models ([50, 44, 43]. The determination of input/output tables is then more efficient and more rapid.

4. Data assimilation and inverse modelling

Suppose that a model has been chosen (Section 1) and a set of appropriate numerical solvers is available (Section 2). The practical use of the resulting CTM deeply relies on data assimilation techniques.

We refer to [57] for more details and some examples. We will only give here a review of the key topics related to data assimilation in CTMs.

4.1. Background

There are unfortunately many uncertainties in the inputs of CTMs: for instance, initial conditions are poorly known and emission fluxes are given with a low accuracy.

Data assimilation consists in coupling the numerical outputs and the observational data in order to lower uncertainty. We refer for instance to [61].

Let us formulate the previous discretized models with the following input/output description:

$$C_{n+1} = F(C_n, \Psi_n) \tag{37}$$

with C_n the outputs (the concentrations) at time t_n for all the grid cells and Ψ_n the vector of forcing parameters (for instance boundary conditions, meteorological data, physical parameters). F stands for the discretized operator mapping time t_n to time t_{n+1}.

Apart from the numerical model, we have at our disposal some observational data from heterogeneous nature, provided by terrestrial networks (Air Quality Monitoring Networks), airborne data or satellital data. If C stands for the chemical state of the atmosphere, we will write $H(C)$ the available observations of C, where H is the observation operator. The measured data are then:

$$obs = H(C) + \epsilon^{(o)} \tag{38}$$

with $\epsilon^{(o)}$ the observation error, supposed to be given by a random Gaussian process $N(0, \Sigma^{(o)})$ with $\Sigma^{(o)}$ the observation error covariance matrix.

To couple numerical outputs and observational data, a cost function $J(u)$, which describes the residual errors between observations and models, is minimized with respect to some uncertain parameters u. This function is typically expressed as:

$$J(u) = (r^{(o)}, (\Sigma^{(o)})^{-1} r^{(o)}) + (r^{(b)}, (\Sigma^{(b)})^{-1} r^{(b)}) \tag{39}$$

with $r^{(o)} = obs - H(C(u))$ the residual observation error and $r^{(b)} = u - u_b$ the residual background error. $(.,.)$ is the usual Euclidean scalar product. $C(u)$ is the model output computed with u. u_b is a background value (often a climatological value) describing the a priori knowledge, with a background error covariance matrix $\Sigma^{(b)}$.

Data assimilation is used to overcome the following problems:

- *Forecast:* $u = C_0$ ([14]).

  Initial conditions are computed in order to minimize forecast errors.

- *Inverse modelling of emissions:* $u = E$ (with previous notations).

  The objective is to improve emission inventories performed by a bottom-up procedure. Forecast may be an objective as well.

- *Inverse modelling of physical parameters:* $u = k$ (for instance kinetic rates).

  Observation are then typically given by smog chamber experiments (on the contrary to the previous cases, field data). The objective is to improve the accuracy for poorly known physical data.

Many algorithms are used in order to solve the data assimilation problem:

- Sequential methods are related to estimation theory and Kalman Filtering ([6]). They have been extensively used, especially for stratospheric chemistry ([24]).

- Variational method (following [28] in meteorology) are related to an iterative minimizing procedure, related to gradient algorithms.

  ∇J has to be computed, which may be unaffordable for a large number of control parameters. This leads to the use of the so-called *adjoint* models, to be derived from the initial forward models. We

refer to [57] for a presentation of some computational aspects related to adjoint CTMs. Modern automatic differentiation tools ([40] for instance) may be powerful techniques in this framework.

Notice that the wide-range of characteristic times for chemical kinetics (see Subsection 3.4.) may impact the quality of data assimilation for short-lived species ([3, 42])

4.2. Some numerical aspects

The development of an adjoint model of a comprehensive 3D CTM may be a difficult task. Consider a cost function $J = 1/2 \sum_{i=1}^{i=n} ||obs_i - HC_i||^2$. In practice, a CTM may be algorithmically described as follows

1. Initialization of time-independent data,
2. Time loop (labeled by $1 \leq i \leq n$):
   - read forcing data Ψ_{i-1} (this contains all the data needed by the solver for computing concentrations at time t_i),
   - compute new state $C_i = F(C_{i-1}, \Psi_{i-1})$,
   - compute the cost function at time t_i: $J_i = 1/2||obs_i - HC_i||^2$,
   - update cost function: $J = J + J_i$.

For the specific case of data assimilation of initial condition ($u = C_0$), the above formula is:

$$\begin{aligned} \nabla_u J &= \sum_{1 \leq i \leq n} (\frac{\partial C_1}{\partial C_0})^T_{|t_0} \cdots (\frac{\partial C_i}{\partial C_{i-1}})^T_{|t_{i-1}} (\frac{\partial J_i}{\partial C_i})^T_{|t_i} \\ &= \ldots \times [(\frac{\partial J_{n-2}}{\partial C_{n-2}})^T + (\frac{\partial C_{n-1}}{\partial C_{n-2}})^T [(\frac{\partial J_{n-1}}{\partial C_{n-1}})^T + (\frac{\partial C_n}{\partial C_{n-1}})^T (\frac{\partial J_n}{\partial C_n})^T]] \end{aligned} \tag{40}$$

and the algorithm now reads:

1. $\hat{C}_n = (\frac{\partial J_n}{\partial C_n})^T$,
2. backward time loop (labelled by $n \geq i \geq 1$):
   - read forcing data Ψ_{i-1},
   - read saved state C_{i-1},
   - compute $\hat{C}_{i-1} = (\frac{\partial C_i}{\partial C_{i-1}})^T_{|C_{i-1}, \Psi_{i-1}} \hat{C}_i$,
   - update $\hat{C}_{i-1} = \hat{C}_{i-1} + (\frac{\partial J_{i-1}}{\partial C_{i-1}})^T$.
3. the output is $\nabla_{C_0} J = \hat{C}_0$.

An adjoint model will compute $z \rightarrow (\frac{\partial C_i}{\partial C_{i-1}})^T_{|C_{i-1}, \Psi_{i-1}} z$. We refer to the next section for an application to a CTM.

5. The example of POLAIR: a 3D Eulerian model for atmospheric dispersion

5.1. Description of POLAIR3D

POLAIR3D is a three-dimensional chemistry-transport model (CTM) developed at ENPC/Air Team for modelling of transboundary dispersion (e.g. acid rains) and forecast of regional photochemical events.

It solves Eq. (1) with state-of-the-art numerical solvers as described in Section 2. The numerical schemes used are a third-order direct space time scheme with Koren-Sweby flux limiting for advection ([66])), the second order Rosenbrock solver ([13]) for chemistry and diffusion and a classical three point scheme for spatial discretization of diffusion. "Strang" splitting operator or AMF may be used to solve the model equation. Emission and dry deposition are treated as lower boundary condition of vertical diffusion. Forcing by a larger-scale model is implemented as boundary condition of advection.

A key point for a CTM is the ability to represent sophisticated chemical mechanisms. In practice, a symbolic preprocessor for multiphase chemistry, called SPACK (Simplified Preprocessor for Atmospheric Chemical kinetics, [13]) is used to generate chemical models (e.g. MOCA [2], CBM4 [16] or RADM2 [59]). The input of SPACK is a syntaxic description of the chemical mechanism. For instance:

```
% GAS PHASE REACTION G3
OH + O3 > HO2
KIN ARR2 1.6d-12  -940.
%
%AQUEOUS PHASE REACTION A2
AO3 + O*2- > AOH + OH- + 2.  AO2
KIN ARRC2 1.5D9 -1500.0D0 298.
%
% ESTABLISHED HENRY'S EQUILIBRIUM FOR O3
O3 =H= AO3
KIN ARRC2 1.1D-2  2300.0D0  298.
```

In this mechanism there is one gas-phase reaction (G3), one liquid-phase reaction (A2) and in-cloud scavenging for ozone. The output of SPACK is a set of Fortran files used in the time integration of the chemical mechanism. SPACK may be directly coupled to ODYSSEE ([40]) in order to generate automatically the related adjoint models.

We refer to [11] for an application to sensitivity analysis of a comprehensive multiphase model.

5.2. Numerical platform

POLAIR3D has been developed in order to be fully modular as far as applications are concerned, e.g.,

- acid rain computations at the continental scale have already been performed by comparison to the Diffeul model carried on at Electricité de France ([69]),
- dispersion of radionucleides over Europe and inverse modelling of sources in the framework of ETEX2 have been tested,
- simulation of photochemical events (for instance, comparison with results of the field campaign ESQUIF over Paris.

Some typical outputs can be found in figure 4 for the continental scale (the grid cell is typically one degree latitude $\times$ one degree longitude) and in figure 5 for the urban scale (the grid cell is 6 km $\times$ 6 km, the emission data are provided by AIRPARIF, the Air Quality Board for Paris).

Moreover, the tangent linear and adjoint versions of POLAIR3D are available. They have been developed by automatic differentiation (with ODYSSEE, [40]) in the framework of an INRIA cooperative research action, Comode (http://www.enpc.fr/cereve/HomePages/sportiss/polaire/pole/projets/Ecomode.html).

The speed-up ratio (CPU of adjoint version/CPU of the forward version) is of magnitude 5, which is a result comparable to the theoretical attended ratio.

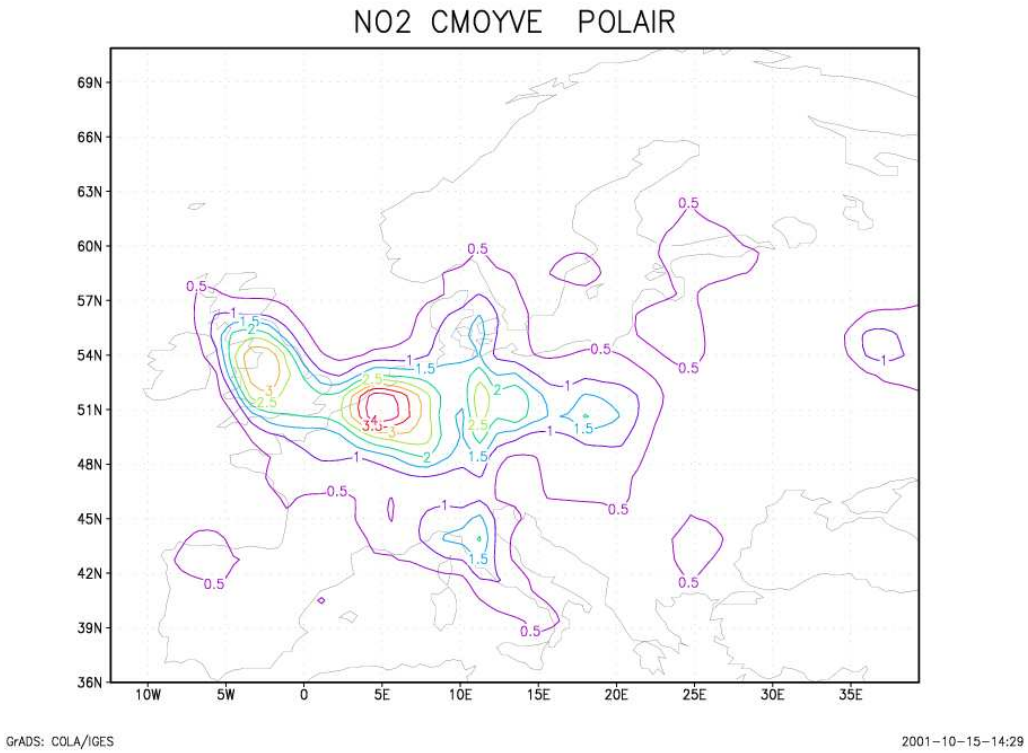

Figure 4. Average NO_2 over vertical columns for 1997

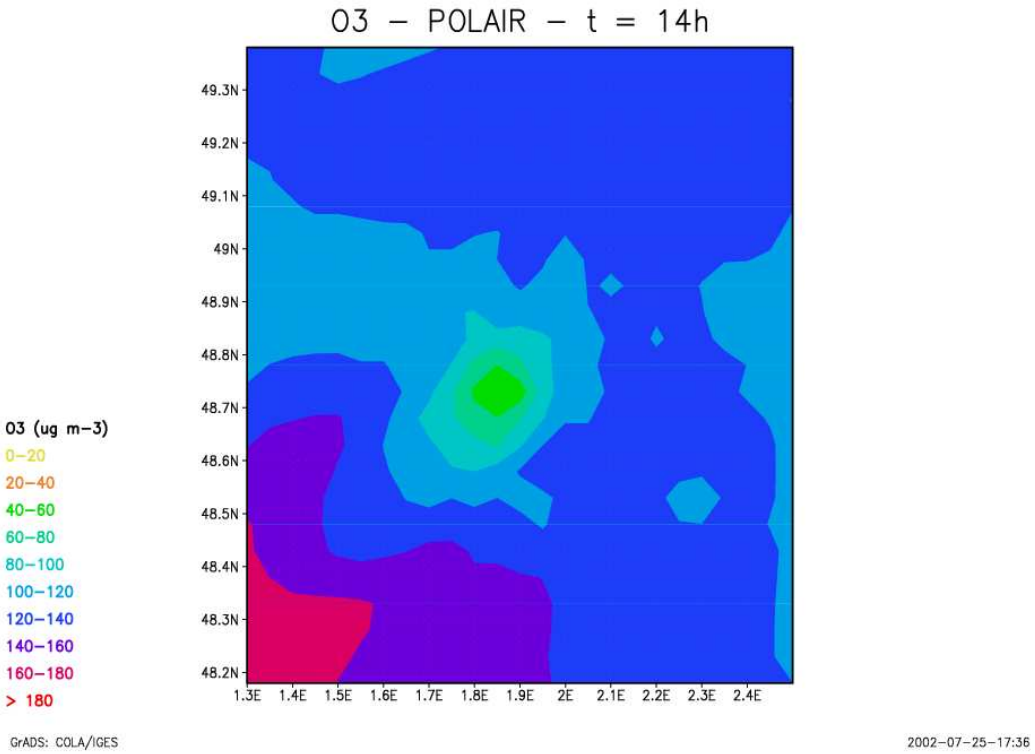

Figure 5. O_3 computed by POLAIR for the 18/7/1999 over Paris

5.3. Future extension

POLAIR3D is a research numerical platform for Air Quality modelling at regional scales. The current extension concern multiphase models (aerosols) with size-resolved and modal approaches. Many applications are concerned ranging from pesticides dispersion to heavy metals dispersion.

POLAIR3D will be inserted in a modelling chain with appropriate post-processing tools (for instance, use of uncertainties propagation tools for impact studies, [62]).

Some open issues

We have briefly reviewed in this paper some of the main issues for air pollution modelling. What will be the next generation of CTMs ?

More accurate parameterizations (for instance: chemistry/dynamics turbulence in the Boundary Layer) will probably be used.

The growing extension to multiphase models has many numerical consequences and developing modular and fast codes describing size-resolved aerosol models remains a challenging point. Many physical aspects remain poorly known (for instance nucleation, transition between dry and wet aerosols).

Coupling of current comprehensive CTMs with radiative transfer models is another key issue for the study of the climate forcing due to aerosols: the resulting coupled models will increase the computational burden. Another application is the direct assimilation of radiative data.

Data assimilation and inverse modelling through 4D-var methods are more and more widespread tools and lead to a dramatic increase of CPU costs. Optimal design of the related environmental monitoring network ([56]) is probably one of the next points to be investigated and needs even more computational resources. Ensemble prevision ([35] could be an issue for air quality forecast.

More generally, the resulting models will be inserted in a modelling chain: ranging from emission models to impact models, with various purposes (forecast, inverse modelling, sensitivity analysis, uncertainty propagation).

These points (among others) motivates the need for a continued development of the current Air Pollution Models with state-of-the-art multiphase models, numerical algorithms and modelling tools.

References

[1] Ackermann, I. J., Hass, H., Memmesheimer M., Ziegenbein, C. and Ebel, A. (1995). The parameterization of the sulfate-nitrate-ammonia aerosol system in eurad. *Meteorol.Atmos.Phys.*, **57**, 101–114.

[2] Aumont, B. (1994). *Modélisation de la chimie de la basse troposphère continentale: développement et tests d'un modèle chimique condensé*. PhD thesis, Université Paris VII.

[3] Austin, J. (1992). Towards the four dimensional assimilation of stratospheric chemical constituents. *J.Geo.Res.*, **97**(D2), 2569–2588.

[4] Berkvens, P. J. F., Botchev, M. A., Krol, M. C., Peters, W. and Verwer, J. G. (2002). *Solving vertical transport and chemistry in air pollution models*, chapter 4. Springer Verlag.

[5] Binkowski, F. S., and U. Shankar, U. (1995). The regional particulate matter model : Model description and preliminary results. *Journal of geophysical research*, **100**(26), 191–209.

[6] Cohn, S. E. (1997). An introduction to estimation theory. *J. Meteor. Soc.Japan*, **75**, 257–288.

[7] Conway, E., Hoff, D. and Smoller, J. (1978). Large-time behaviour of solutions of systems of nonlinear reaction-diffusion equations. *SIAM J. Appl. Math.*, **35**(1), 1–16.

[8] Debry, E., Jourdain, B. and Sportisse, B. (2001). Modelling of aerosol dynamics: a stochastic algorithm. In B. Sportisse, editor, *APMS 2001*, Geosciences, 308–319. Springer.

[9] Debry, E. and Sportisse, B. (2001). Un point sur la simulation numérique des aérosols atmosphériques. In *Congrès de l'ASFERA. Pollution Atmosphèrique*, 74–80.

[10] Dekker, K. and Verwer, J. G. (1984). *Stability of Runge-Kutta methods for stiff non linear differential equations*. North Holland.

[11] Djouad, R., Audiffren, N. and Sportisse, B. (2001). Sensitivity analysis using automatic differentiation applied to a multiphase chemical mechanism. Submitted to *Atmospheric Environment*.

[12] Djouad, R. and Sportisse, B. (2001). Use of proper orthogonal decompositions for the reduction of atmospheric chemistry. Submitted to *J. Geophys. Res.*

[13] Djouad, R., Sportisse, B. and Audiffren, N. (2002). Numerical simulation of aqueous-phase atmospheric models: use of a non-autonomous rosenbrock method. *Atmos. Environ.*, **36**, 873–879.

[14] Elbern, H., Schmidt, H. and Ebel, A. (1997). Variational data assimilation for tropospheric chemistry modelling. *J. Geophys. Res.* **102**, 15967–15985.

[15] Gelbard, F., Tambour, Y. and Seinfeld, J. H. (1980). Sectional representations for simulating aerosol dynamics. *Journal of collod and Interface Science*, **76**(2), 541–556.

[16] Gery, M. W., Whitten, G. Z., Killus, J. K. and Dodge, M. C. (1989). A photochemical kinetics mechanism for urban and regional scale computer modelling. *J. Geophys. Research*, **94**(D10), 12925–12956.

[17] Hale, J. K. and Rocha, C. (1987). Varying boundary conditions with large diffusivity. *J. Math. pures et appl.*, **66**, 139–158.

[18] Herrmann, H., Ervens, B., Jacobi, H. W., Nowacki, P., Wolke, R. and Zellner, R. (2000). Capram 2.3: a chemical aqueous phase radical mechanism for tropospheric chemistry. *J. Atmos. Chem.*, **36**, 231–284.

[19] Hesstvedt, E., Hov, O. and Isaksen, I. S. A. (1978). Qssas in air polution modelling: comparison of two numerical schemes for oxydant prediction. *Int. J. Chem. Kinet.*, **10**, 971–994.

[20] Hundsdorfer, W. H. (1996). *Numerical solution of advection-diffusion-reaction equations*. Technical Report NM-N9603, CWI.

[21] Jacobson, M. Z., Turco, R.P., Jensen, E. J. Toon, O. B. (1994). modelling coagulation among particles of different composition and size. *Atmospheric Environment*, **28**(7), 1327–1338.

[22] Jaecker-Voirol, A. and Mirabel, P. (1989). Heteromolecular nucleation in the sulfuric acid-water system. *Atmos.Environ.*, **23**, 2053–2057.

[23] Karamchandani, P. and al. (2001). *Development and application of a state-of-the-science plume-in-grid model*. Technical report, AER Inc.

[24] Khattatov, B. and al. (1999). Assimilation of photochemically active species and a case analysis of uars data. *J.Geophys. Res.*, **104**, 18715–18737.

[25] Kunisch, K. and Volkwein, S. (2001). Galerkin proper orthogonal decomposition methods for parabolic problems. To appear in *Numer. Math.*

[26] Lam, S. H. and Goussis, D. A. (1994). The csp method for simplifying kinetics. *Int.J. Chem. Kinet.*, **26**. 462–484.

[27] Lanser, L. and Verwer, J. G. (1998). Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling. In *Proceedings 2nd Meeting on Numerical methods for differential equations*. Coimbra, Portugal, February.

[28] Le Dimet, F. X. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus*, **38**A, 97–110.

[29] Lelieveld, J. and Crutzen, P. J. (1991). The role of clouds in tropospheric chemistry. *Journal of Atmospheric Chemistry*, **12**, 229–267.

[30] Louis, J. L. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Met.*, **17**, 197–202.

[31] Madronich, S. (1987). Photodissocation in the atmopshere. 1. actinic flux and the effects of ground reflections and clouds. *J. Geophys. Res.*, **92**, 9740–9752.

[32] Marchuk, G. I. (1986). *Mathematical models in environmental problems*, North Holland.

[33] Maas, U. (1993). *Automatische Reduktion von Reaktionsmechanismen zur Simulation reaktiver Stromungen*. PhD thesis, Stuttgart Univers.

[34] McRae, G. J., Goodin, W. R. and Seinfeld, J. H. (1982). Numerical solution of the atmospheric diffusion equation for chemically reacting flows. *J.Comp. Phys.*, **45**, 1–42.

[35] Miller, R. N., Carter, E. F. Jr and Blue, S. T. (1998). *Data assimilation into nonlinear stochastic models*. Technical report, Oregon State Univ.

[36] Molemaker, M. J. and Vila-Guerau de Arellano, J. (1998) Control of chemical reactions by convective turbulence in the boundary layer. *J. of the Atm. Sc.*, **55**, 568–579.

[37] Morita, Y., Ninomiya, H. and Yanagida, E. (1994). Nonlinear pertubation of boundary values for reaction-diffusion systems: inertial manifolds and their applications. *SIAM J. Math. Anal.*, **25**(5), 1320–1356.

[38] Nenes, A., Pandis, S. N. and Pilinis, C. (1998). Isorropia : A new thermodynamic equilibrium model for multi-component inorganic aerosols. *Aquatic geochemistry*, **4**, 123–152.

[39] Pandis, S. N., Wexler, A. S. and Seinfeld, J. H. (1993). Secondary organic aerosol formation and transport -ii. predicting the ambient secondary organic aerosol size distribution. *Atmospheric Environment*, **27**A(15), 2403–2416.

[40] Papegay, Y. and Faure, C. (1997). *Odyssé version 1.6: The language reference manual*. Technical Report RT-211, INRIA.

[41] Pilinis, C. (1990). Derivation and numerical solution of the species mass distribution equations for multicomponent particulate systems. *Atmospheric Environment*, **24**A(7), 1923–1928.

[42] Quélo, D., Sportisse, B., Berroir, J.P. and Charpentier, I. (2002). Some remarks concerning inverse modelling and data assimilation for slow-fast atmospheric chemical kinetics. In B. Sportisse, editor, *APMS 2001*, Geosciences, 499–513. Springer.

[43] Rabitz, H. and Alis, O. F. (1999). General foundations of high-dimensional model representations. *J. Math. Chem.*, **25**, 197–223.

[44] Rabitz, H., Alis, O. F., Shorter, J. A. and K. Shim, K. (1999). Efficient input-output model representations. *Journal of Computer Physics Communication*, **117**, 11–20.

[45] Sandu, A. (2001). A spectral method for solving aerosol dynamics. Submitted to *Applied Numerical Mathematics*.

[46] Sandu, A., Verwer, J. G., Van Loon, M., Carmichael, G., Potra, F. A., Dabdub, D. and J.H. Seinfeld, J. H. (1997). Benchmarking stiff odes solvers for atmospheric chemistry problems i: implicit versus explicit. *Atmos. Environ.*, **31**, 3151–3166.

[47] Sanz-Serna, J. M. (1997). *The State of the art in numerical analysis*, chapter Geometric integration, pages 121–143. Clarendon Press, Oxford.

[48] Schwartz, S. E. (1986). *Mass-transport considerations pertinent to aqueous phase reactions of gases in liquid-water clouds*. NATO ASI Series, Vol. G6, Springer Verlag, Chemistry of multiphase atmospheric systems.

[49] Seinfeld, J. H. (1985). *Atmospheric physics and chemistry of Air Pollution*. Wiley.

[50] Shorter, J. A. and Precila C.Ip. (1999). An efficient chemical kinetics solver using high dimensional model representation. *Journal of Physical Chemistry*, **103**, 7192–7198.

[51] Sportisse, B. (2000). An analysis of operator splitting techniques in the stiff case. *J. Comp. Phys.*, **161**, 140–168.

[52] Sportisse, B. (2001). Box models versus eulerian models in air pollution modelling. *Atm. Environ.*, **35**, 173–178.

[53] Sportisse, B. and Djouad, R. (2000). Reduction of chemical kinetics in air pollution modelling. *J. Comp. Phys.*, **164**, 354–376.

[54] Sportisse, B. and Djouad, R. (2002). Mathematical investigation of mass transfer for atmospheric pollutants into a fixed droplet with aqueous chemistry. Accepted for publication in *Journal of Geophysical Research Atmospheres*.

[55] Sportisse, B. and Du Bois, L. (2002) Numerical and theoretical investigation of a simplified model for the parameterization of below-cloud scavenging by falling raindrops. *Atmos.Environ.*, **36**, 5719–5727.

[56] Sportisse, B. and Quélo, D. (2002). Observational network design, adaptive observations and targeting strategies in atmospheric data assimilation: sketch of a methodological review. In *INRIA-CEA-EDF Summer School on Data Assimilation for Geophysical flows*.

[57] Sportisse, B. and Quélo, D. (2002) Data assimilation and inverse modelling of atmospheric chemistry. *Special issue of the Journal of Indian National Science Academy on "Advances in Atmospheric and Oceanic Sciences"*. To be published.

[58] Srivastava, R. K., Mc Rae, D. S. and Odman, M. T. (2001). Simulation of dispersion of a power plant plume using an adaptive grid algorithm. *Atmos.Environ.*, **35**, 4801–4818.

[59] Stockwell, W. R., Middleton, P., Chang, J. and Tang, X. (1990). The second regional acid deposition model chemical mechanism for regional air quality modelling. *J.of Geophysical Research*, **95**(D10), 16343–16367.

[60] Strang, G. (1968). On the construction and comparison of difference schemes. *SIAM J. Numer. Anal.*, **5**, 506–517.

[61] Tarantola, A. (1987). *Inverse Problem Theory*. Elsevier Amsterdam.

[62] Tatang, M. A. (1995). *PhD thesis*. PhD thesis, MIT.

[63] Tiedke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large scale models. *Monthly Weather Review*, **117**, 1779–1800.

[64] Tomlin, A., Berzins, M., Ware, J.,Smith, J. and Pilling, M. J. (1997). On the use of adaptative gridding methods for modelling chemical transport from multi-scales sources. *Atm. Env.*, **31**(18), 2945–2959.

[65] Vasileva, A. B., Butuzov, V. F. and Kalachev, L. V. (1994). *The boundary function method for singular perturbation problems*. SIAM monographs.

[66] Verwer, J. G., Hundsdorfer, W. H., and J.G. Blom, J. G. (1998). Numerical time integration for air pollution models. In *Proceedings of the Conference APMS'98*. ENPC-INRIA, 26–29.

[67] Verwer, J. H., Spee, E. J., Blom, J. G. and Hundsdorfer, W. H. (1999). A second order rosenbrock method applied to photochemical dispersion problem. *SIAM J. SCI. COMPUT.*, **20**(4), 1456–1480.

[68] Verwer, J. G. and Sportisse, B. (1998). *A note on operator splitting analysis in the stiff linear case*. Technical Report MAS-R9830, CWI.

[69] Wendum, D. (1998). Three long range transport models compared to the etex experiments: a performance study. *Atmos. Environ.*, **32**, 4297–4305.

[70] Wesely, M. L. (1989). Parameterization of surface resistance to gaseous-dry deposition in regional scale numerical models. *Atmos. Environ.*, **23**, 1293–1304.

[71] Wexler, A. S. and Seinfeld, J. H. (1990). The distribution of ammonium salts among a size and composition dispersed aerosol. *Atmospheric Environment*, **24**A(5), 1231–1246.

[72] Whitby, E. R. and McMurry, P. H. (1997). Modal aerosol dynamics modelling. *Aerosol Science and Technology*, **27**, 673–688.

B. Sportisse, J. Boutahar, E. Debry, D. Quélo & K. Sartelet
Centre d'Enseignement et de Recherche sur l'Eau, la ville et l'environnement,
'Ecole Nationale des Ponts et Chaussées (ENPC-CEREVE), Air Team,
rue Blaise Pascal, 77455 Champs sur Marne, France
sportisse@cereve.enpc.fr