Comput Geosci (2007) 11:159–181
DOI 10.1007/s10596-006-9036-4

REVIEW PAPER

A review of current issues in air pollution modeling and simulation

Bruno Sportisse

Received: 20 November 2006 / Accepted: 4 December 2006 / Published online: 19 January 2007
© Springer Science + Business Media B.V. 2007

Abstract Air pollution modeling is now a mature field, and comprehensive numerical models (the chemistry-transport models) are used in many applications. This article aims at reviewing the main issues from the point of view of applied mathematics and computational physics (as viewed by the author). We address topics such as subgrid parameterization, numerical algorithms with a focus on aerosol simulation, data assimilation and inverse modeling, reduction of high-dimensional models and propagation of uncertainties. Even if this article is strictly related to air pollution modeling, many issues and methods can be extended to dispersion of tracers in other media (for instance, water).

Keywords air pollution modeling · chemistry-transport models

1 Introduction

Air pollution modeling is related to many other fields, ranging from the study of short-range dispersion of species (typically accidental release in the case of an industrial hazard) to atmospheric chemistry and climate change [118]. Many scales are concerned: local scale (accidental release), regional scale (photochemistry, urban pollution), continental scales (transboundary pollution with the example of acid rains), global scale (atmospheric chemistry in the stratosphere, oxidizing power of the troposphere, etc). Many species are of interest: ozone and volatile organic compounds (photochemistry), trace metals, mercury, methane, carbon monoxide, particulate matter (aerosols), radionuclides, biological species, etc. There are also many applications of the resulting numerical models: understanding of physical processes (assessment of the impact of a given process), environmental forecast (performed by an emergency center, for instance), impact studies (of emission sources), sensitivity analysis (with respect to different scenarios of emission), inverse modeling (of uncertain emissions), etc.

All these applications are usually based on the so-called chemistry-transport models (CTM). The objective of a CTM is to simulate the time evolution of spatial fields for a set of chemical species (by extent of radionuclides and biological species).

We refer to Fig. 1 for the processes to be described. Some *primary* species (for instance, nitrogen oxides [NO_x] or volatile organic compounds [VOC]) are emitted either by anthropogenic or biogenic sources (surface emission or point emission). The species are then vertically diffused in the atmospheric boundary layer by turbulent eddies related to both mechanical forces (wind shear) and thermal forces (buoyancy). The horizontal motion is related to wind advection. Gas-phase chemical reactions, mainly related to the oxidizing power of the atmosphere and to the radiative fluxes (through photolysis), lead to the production of *secondary* species (typically ozone, O_3). Mass transfer between gas phase, aqueous phase (cloud droplets) and particulate matter (solid or liquid particles in suspen-

---

This article is based on a plenary talk given at SIAM Geosciences, Avignon, June 2005.

B. Sportisse (✉)
CEREA, Joint Laboratory ENPC/EDF R&D,
CLIME project ENPC/INRIA, Paris, France
e-mail: sportiss@cerea.enpc.fr

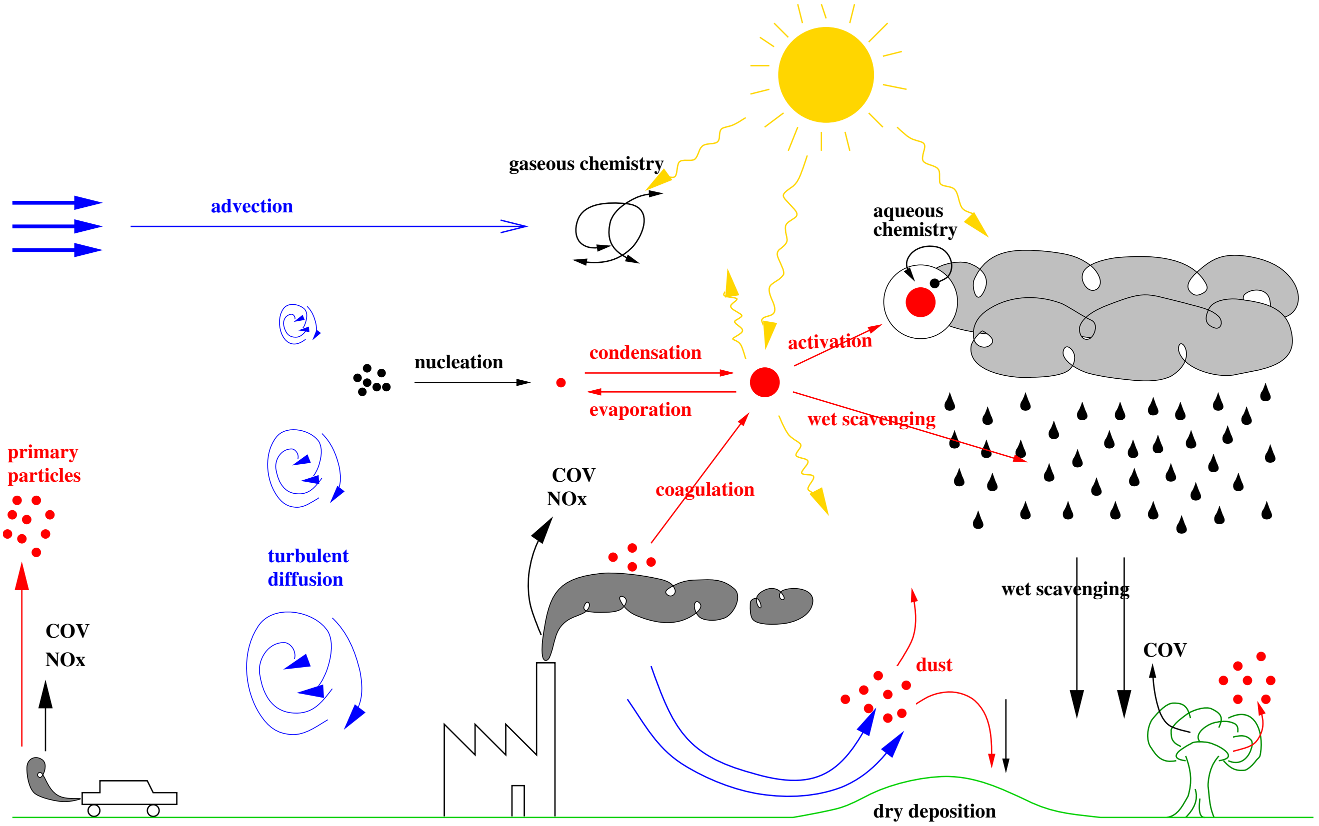

Fig. 1 Chemistry-transport models

sion) may also happen. Moreover, the evolution of aerosols is governed by microphysical processes such as nucleation (the formation of small clusters of gaseous molecules), coagulation (collision between particles), condensation/evaporation (mass transfer between the semi-volatile species and particles) or activation (the growth of aerosols to cloud droplets through condensation of water vapor). The loss processes from the atmosphere are dry deposition (when gases or particles impinges upon and stick to the surface) and wet scavenging (washout by rains).

The underlying model is the equation of atmospheric dispersion. For instance, the evolution of the gas-phase concentration c_i (typically in μg m^{-3}) is governed by the following equation [118]:

$$\frac{\partial c_i}{\partial t} + \underbrace{\mathrm{div}\,(Vc_i)}_{\text{advection}} = \underbrace{\mathrm{div}\left(\rho K \nabla \frac{c_i}{\rho}\right)}_{\text{turbulent diffusion}} + \underbrace{\chi_i(c,t)}_{\text{gas-phase chemistry}} - \underbrace{\Lambda_i(x,t)c_i}_{\text{scavenging}} + \underbrace{S_i(x,t)}_{\text{volume sources}} \qquad (1)$$

$V=(u,v,w)$ stands for the wind velocity (the components are given for the directions x, y and z), ρ for the air density and K for the eddy diffusivity matrix (in practice, a diagonal matrix); χ_i is the ith component of the chemical production (given by a chemical mechanism that describes the reactions among a set of chemical reactants); the scavenging term is a parameterization of the scavenging processes by cloud droplets (rainout or in-cloud scavenging) and by rain drops (washout or below-cloud scavenging); the volume source is related to point sources (typically a chimney).

Mathematically speaking, this is a system of advection–diffusion–reaction partial differential equations (PDEs), coupled through the chemical production term. The initial conditions and the boundary conditions have to be added, for instance at ground:

$$-K\frac{\partial c_i}{\partial z} = \underbrace{E_i(x,t)}_{\text{surfacic emissions}} - \underbrace{v_i^{\mathrm{dep}}(x,t)c_i}_{\text{dry deposition}} \qquad (2)$$

E_i is the flux of surface emissions (typically related to biogenic emissions and to traffic emissions); v_i^{dep} is the dry deposition velocity, which strongly depends on the land use coverage. z is the vertical coordinate.

In the last two decades, many efforts have been devoted to the numerical simulation of these equations. The first review paper is [82]. A general overview of numerical analysis for the CTM is updated in [149]. We can also refer to [170].

In recent years, the maturity of CTM has given raise to many new computational and methodological issues that are not only related to numerical analysis. The aim of this paper is to try to review these issues.

This paper is structured as follows. The parameterizations are the key components of a CTM. In Section 2, we review some issues related to subgrid parameterizations, with a focus on methodological issues. Once the model is defined, the numerical simulation may be a challenging point. In Section 3, we address some classical numerical difficulties (splitting, stiffness). In Section 4, we focus on the most difficult part, the numerical simulation of the general dynamics equation for aerosols. These first three sections are related to the so-called "forward model". Sophisticated uses of CTM (for instance, for environmental forecast) generate other issues such as inverse modeling and data assimilation (Section 5) or the evaluation of uncertainties (Section 6). In many applications (for instance, for integrated modeling), comprehensive models are not useable as such and the search for proxy/reduced models is still an issue (Section 7). As far as implementation is concerned, modern architectures for CTM are also required (Section 8).

As far as notations are concerned, we have tried to use similar notations throughout the article. Because of existing standard notations for some topics, this has unfortunately not be always possible. The hope is that it does not perturb the reader.

2 Subgrid parameterization

Subgrid parameterization is a key issue in air pollution modeling. Indeed, many physical processes occur at scales much smaller than the numerical scales (typically given by the mesh dimension).

In the following, $\langle \Psi \rangle$ is the averaged field associated to Ψ (to be defined in a rigorous way by the averaging framework chosen for the turbulent model). In practice, this corresponds to the computed field (at the scale of the numerical model). The fluctuation $\Psi^{'}$ is defined as $\Psi^{'} = \Psi - \langle \Psi \rangle$.

For any nonlinear process that implies terms such as $\Psi_1 \Psi_2$, a correlation term $\left\langle \Psi_1^{'} \Psi_2^{'} \right\rangle$ has to be parameterized as a function of the averaged fields $\langle \Psi_1 \rangle$ and $\langle \Psi_2 \rangle$. This is the classical *closure* problem of the unresolved scales.

We briefly illustrate this issue with four examples: eddy diffusion, microphysical mass transfer, coupling between turbulence and chemistry and plume-in-grid modeling.

2.1 Eddy diffusion

A first example is given by eddy diffusion. In Eq. 1, K stands for the eddy diffusion matrix. Molecular diffusion is neglected in the atmospheric boundary layer (apart from a very thin laminar layer just above the ground), and diffusion is related to turbulent fluxes. After having averaged the dispersion equation written for the "true" fields, the advective terms generate an advective mean term, $\mathrm{div}\left(\langle V \rangle \langle c \rangle\right)$, and a correlation term, $\mathrm{div}\left(\left\langle V^{'} c^{'} \right\rangle\right)$. With the K-theory [70], the correlation between the vertical velocity w and the concentration c is usually expressed as:

$$\left\langle w^{'} c^{'} \right\rangle = - \langle \rho \rangle \, K_z \frac{\partial}{\partial z} \left(\frac{\langle c \rangle}{\langle \rho \rangle} \right) \tag{3}$$

We advocate the above formulation rather than $\left\langle w^{'} c^{'} \right\rangle = -K \nabla \left(\langle c \rangle\right)$ to get 0 if $\frac{\langle c \rangle}{\langle \rho \rangle}$ is constant. With this formulation of the correlation term, one has therefore only to check mass consistency for the advection equation for averaged fields (see Section 3.3 and [131]).

The vertical eddy coefficient K_z has then to be parameterized on the basis of available meteorological fields (for instance, [74] or [141]). This is a key component of any CTM. Notice that the horizontal eddy diffusion has also to be parameterized. In many models, horizontal diffusion is neglected because of the lack of well-founded models and of the magnitude of numerical diffusion induced by advection solvers.

2.2 Scavenging

Apart from turbulence, subgrid parameterization is needed for microphysical processes related to mass transfer among gas phase and condensed matter (cloud droplets, rain droplets, aerosols). Let us recall the typical microphysical scales: the aerosol diameter ranges from a few nanometers to, let say, 10 μm; the diameter of a cloud droplet ranges from a few micrometers to, let say, 100 μm; the diameter of a raindrop is of magnitude 0.1-0.5 mm. These scales have to be compared to the numerical scales. At local scale, the horizontal magnitude of a grid cell is typically a few meters, whereas it is a few kilometers at regional scale and a few tens of kilometers at continental scale.

2.2.1 In-cloud scavenging

Mass transfer and aqueous-phase chemistry occur inside cloud droplets, and this constitutes a key process for acid rains and sulfate formation for aerosols. The underlying models are diphasic models that describe the time evolution of gas-phase (interstitial) concentrations (c_g) and aqueous-phase concentrations (c_a).

If Ω_a stands for the volume defined by a cloud droplet (supposed to be a sphere of surface Γ and of radius a) and Ω_g stands for the interstitial gas phase, the microphysical model is [118]:

$$\begin{cases} \dfrac{\partial c_g}{\partial t} + D_g \Delta c_g = \chi_g(c_g) \text{ over } \Omega_g \\ \dfrac{\partial c_a}{\partial t} + D_a \Delta c_a = \chi_a(c_a) \text{ over } \Omega_a \end{cases} \tag{4}$$

The mass transfer is given at the droplet surface Γ by:

$$D_g \nabla c_g \cdot \vec{n} = -D_a \nabla c_a \cdot \vec{n} = \frac{1}{4}\alpha \bar{v} \left(\frac{c_a}{HRT} - c_g \right) \tag{5}$$

χ_g (χ_a) stands for the gas-phase (aqueous-phase) production term. D_g (D_a) stands for the gas-phase (aqueous-phase) molecular diffusivity. $\vec{n}$ is the unitary normal vector to Γ (outward oriented); α is an accommodation coefficient, $\bar{v}$ is the thermal velocity for the chemical species, H is the Henry's coefficient (which describes the partitioning between the gas phase and the aqueous phase for a given species), R is the gas constant and T is the temperature.

It is of course not tractable to solve these local PDEs coupled to the CTMs. A key property is that there is a wide range of timescales: molecular diffusion is much faster than chemistry at these scales. In practice, one uses a *lumped parameter assumption* by replacing the microphysical concentrations c_a and c_g by averaged concentrations, for instance, $\bar{c}_a = \frac{1}{\Omega_a} \int_{\Omega_a} c_a(\omega, t) \mathrm{d}\omega$. The averaged concentrations are then given by a system of ordinary differential equations (ODE):

$$\begin{cases} \dfrac{\mathrm{d}\bar{c}_g}{\mathrm{d}t} = \chi_g(\bar{c}_g) - L\, I(\bar{c}_a, \bar{c}_g) \\ \dfrac{\mathrm{d}\bar{c}_a}{\mathrm{d}t} = \chi_a(\bar{c}_a) + I(\bar{c}_a, \bar{c}_g) \end{cases} \tag{6}$$

The mass transfer flux I is parameterized as $I(\bar{c}_a, \bar{c}_g) = k_{mt} \left(\bar{c}_g - \dfrac{\bar{c}_a}{HRT} \right)$. $k_{mt} = (a^2/3D_g + 4\bar{v}/3\alpha)^{-1}$ is the kinetic mass transfer coefficient. L is the liquid water content (defined as the volume of Ω_a over the volume of Ω_g).

This parameterized model may be rigorously derived by asymptotic expansions with respect to a small Peclet–Dahmköhler number. One refers for instance to [127]. The mathematical background may be found in [43] for reaction–diffusion systems and in [144] for the method of boundary layer.

There are still many open issues, for instance, the extension to polydisperse droplets or to reactive film models that are not defined by the interface condition (5).

2.2.2 *Below-cloud scavenging*

A second example is given by *below-cloud* scavenging. In this case, the mass transfer occurs between the gas phase and the falling raindrops (of falling velocity U and of radius a). With the same notations as before, the microphysical model is [118]:

$$\begin{cases} \dfrac{\partial c_g}{\partial t} = -\dfrac{L}{\tau_{\text{transfer}}} \left(c_g - \dfrac{c_a}{HRT} \right) \\ \dfrac{\partial c_a}{\partial t} + U \dfrac{\partial c_a}{\partial z} = \dfrac{1}{\tau_{\text{transfer}}} \left(c_g - \dfrac{c_a}{HRT} \right) \end{cases} \tag{7}$$

where τ_{transfer} is a timescale associated to mass transfer (not detailed here). It is similar to the inverse of a kinetic mass transfer coefficient. By applying the boundary layer method in the case of a small liquid water content or of high solubility, this system may be replaced by the parameterization [118]:

$$\frac{\mathrm{d}c_g}{\mathrm{d}t} = -\Lambda_g c_g \tag{8}$$

where $\Lambda_g(p_0, a, HRT, \tau_{\text{transfer}})$ is the so-called scavenging coefficient, with p_0 the rain intensity ([129]).

The parameterization of scavenging for particles is also particularly difficult (see for instance [124] and references therein).

2.3 Segregation effects

Another example is provided by the coupling between chemistry and turbulence. After having averaged the dispersion equation, the chemical term appears as $\left\langle \chi_i(\langle c \rangle + c') \right\rangle$. One usually assumes that this term is equal to $\chi_i(\langle c \rangle)$, which may be a crude approximation for the fastest reactions. The underlying assumption is indeed that the averaging is quicker than any chemical reactions. We refer for instance to the review [154].

A typical example is given by the so-called segregation effect between an updraft of emitted NO and a downdraft of O_3 (entrainment) in the atmospheric boundary layer. The chemical reaction of O_3 with NO (titration) is:

$$O_3 + NO \rightarrow NO_2 + O_2$$

The associated reaction rate is $\omega = k\, c_{O_3} c_{NO}$ with k the kinetic rate. After averaging, the true averaged reaction rate is $\langle \omega \rangle = k \left(\left\langle c_{O_3} c_{NO} \right\rangle + \left\langle c'_{O_3} c'_{NO} \right\rangle \right)$, where we suppose that the kinetic rate has a constant value (it actually depends on temperature). This can also be written as:

$$\langle \omega \rangle = k \left\langle c_{O_3} c_{NO} \right\rangle (1 + I_s) \qquad (9)$$

with $I_s = \frac{\left\langle c'_{O_3} c'_{NO} \right\rangle}{\left\langle c_{O_3} c_{NO} \right\rangle}$ the *segregation intensity*. Mathematically speaking, $I_s \geq -1$; physically speaking, $I_s \leq 0$ because O_3 is associated to downdrafts, whereas NO is associated to updrafts.

In the "laminar" assumption, I_s is taken equal to 0. A parameterization is then required for I_s [85, 155, 156]. Once more, the key parameters are the so-called Dahmköhler numbers defined as the ratio of a transport timescale to a chemical timescale. We also refer to [123] for a simple estimation of the segregation errors in a 1D case.

To date, there are no available parameterizations of I_s to be used in 3D models with comprehensive chemical mechanisms.

2.4 Plume-in-grid modeling

The last example is provided by the so-called plume-in-grid models. Emissions of point sources in a large grid cell induce a numerical artefact by artificially creating diffusion. One way to reduce this diffusion is to parameterize short-term dispersion of freshly emitted pollutants as an alternative to the use of refined meshes (see Section 3). The parameterizations are usually based on Gaussian models (stationary solutions of the dispersion equation). We refer, for instance, to [67]. An issue is the rigorous coupling between long-range Eulerian models and short-range Gaussian models for reactive flows.

These examples illustrate the need for rigorous approaches to derive the appropriate parameterizations. An alternative approach is based on numerical sampling. In [72], a numerical sampling technique (a Latin hypercube method) is, for instance, used to define the parameterization. A probability density function (PDF) is then computed on the basis of repeated calls to a microphysical function.

3 Numerics

We now assume that the model is fixed. The numerical simulation may provide a large number of challenging numerical points, and we pay attention to four numerical issues: operator splitting (the processes are usually not solved simultaneously), time integration of stiff ODEs (resulting from chemistry or aerosol dynamics), advection (with the specific problem of mass consistency) and grid resolution. The most challenging issues are related to aerosol dynamics and are described in the next section.

A detailed review can be found in [149] for some of these topics (excluding aerosol dynamics). We can also refer to the book [61].

3.1 Splitting methods

Chemistry-transport models describe many processes related to dispersion, microphysics and chemistry. Although all these processes are coupled, operator splitting methods (for instance, [81, 163]) are widely used for solving the dispersion equation. There are at least three motivations for this numerical strategy.

First, models can be used as black boxes and are easy to insert in an existing comprehensive three-dimensional model. Second, appropriate specific numerical methods can be used for each process. Third, coupling dispersion and chemistry may result in a huge computational burden because implicit methods are used for chemistry (because of the dispersion of timescales; see below). In fully coupled methods, the size of the state vector is then typically equal to the product of the number of species by the number of grid cells (for finite differences solvers). The resulting Jacobian matrices are then difficult to invert.

For the sake of clarity, we consider two processes that do not depend on time. The evolution equation is therefore:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = A_1(c) + A_2(c) \;, \quad c(0) = c_0 \qquad (10)$$

A_1 (for instance, diffusion) and A_2 (for instance, chemistry) are the source terms associated to the processes.

Let Δt be the splitting timestep (for a CTM, it typically ranges from 100 to 1,000 s at continental scales). We formally write $\mathcal{L}_{A_1+A_2}(c_0, t)$ the solution of equation (10) at time t.

The simplest method is the first-order splitting: one first solves the process A_1 over $[0, \Delta t]$ and then (after having modified the initial condition) the process A_2 over $[0, \Delta t]$. The solution is then $\mathcal{L}_{A_2}(\mathcal{L}_{A_1}(c_0, \Delta t), \Delta t)$. Notice that we can also reverse the sequence between A_1 and A_2.

The so-called Strang splitting [134] eliminates the lack of symmetry of the previous method. There are three steps in the sequence: A_1 is integrated over $[0, \Delta t/2]$, then A_2 is integrated over $[0, \Delta t]$ and finally A_1 over $[0, \Delta t/2]$. The solution is therefore $\mathcal{L}_{A_1}(\mathcal{L}_{A_2}(\mathcal{L}_{A_1}(c_0, \Delta t/2), \Delta t), \Delta t/2)$.

Operator splitting methods induce a loss of accuracy. This is usually estimated with the computation of the local splitting error in the linear case for $A_i(c) = A_i c$

(A_i is a matrix). In this case, $\mathcal{L}_A(c_0, \Delta t) = e^{A\Delta t}\, c_0$. The splitting error is related to the commutation default: $e^{A_1+A_2} = e^{A_2}e^{A_1}$ if and only if $A_1\, A_2 = A_2\, A_1$.

Asymptotic analysis with respect to $\Delta t \to 0$ prove that the Strang splitting method is of second-order (and that the first-order method is well named!).

Some results are available for the reaction–diffusion–advection PDE with the extension to the nonlinear case by using the Lie formalism [71]. The key results are the following ones (boundary conditions are not taken into account):

- Advection and chemistry commute if chemistry does not depend on space and if the wind field is divergence free.
- Advection and diffusion commute if the wind field and the diffusion matrix do not depend on space.
- Diffusion and chemical kinetics commute if chemical kinetics is linear and does not depend on space.

It is clear that the assumption concerning the independence with respect to time and space is not rigorously met for the eddy diffusivity matrix and the temperature field. This is, however, a rather good *local* assumption. On the contrary, the assumption of linearity for chemistry is clearly wrong (chemical reactions are supposed to describe...reactions between reactants!). However, these results give an indication for the main source of errors related to operator splitting methods in air pollution models: chemistry and vertical diffusion have to be handled with special care [51].

A difficult point (not addressed here) is also related to the way the boundary conditions have to be taken into account (see for instance [121]).

For atmospheric chemistry, the previous results (especially the order analysis) cannot be applied as such. A major point is indeed that one process (for instance, chemistry) contains fast dynamics that quickly reaches an equilibrium (see below for more details and the relation to stiffness). In practice, this means that there exist processes whose timescales τ are much lower than the splitting timestep Δt. The asymptotic analysis $\Delta t \to 0$ is therefore no longer justified. We refer, for instance, to [122, 125, 152] for a deeper understanding. One key result is the impact of the splitting sequence: it is advocated to end the sequence with the stiff processes (chemistry for instance in the case of gas-phase chemistry, mass transfer for aerosol dynamics) to ensure that the system is still along the equilibrium manifold at each step. A second result is related to the so-called *order reduction* of the numerical algorithms. Similar results have been obtained in other fields, for instance, in combustion [164]. See also the discussion for oceanography in [58] or for climate models in [161].

Methods that avoid splitting are also used. Among them, one can cite implicit–explicit methods that are based on a partitioning between the stiff processes (with an implicit treatment) and the nonstiff processes (with an explicit treatment). There are many variations of this method (for instance, [88, 136, 147, 162]).

Let us suppose that A_1 and A_2 are the nonstiff process and the stiff process, respectively. A possible implementation (usually referred as *source splitting*) is to first solve A_1 and then to solve A_2 after having added a correction term $\tilde{A}_1 = \dfrac{\mathcal{L}_{A_1}(c_0, \Delta t) - c_0}{\Delta t}$ *without modifying the initial condition*. The result is then: $\mathcal{L}_{A_2+\tilde{A}_1}(c_0, \Delta t)$. The key advantage is that the integration is always performed in the vicinity of the equilibrium manifold of the stiff process [122].

Another approach is the so-called approximate matrix factorizations (also referred as “internal splitting”; see [7]). The idea is to perform splitting at the algebraic level. Suppose that an implicit algorithm (let say Backward Euler) is used for solving equation (10) in the linear case.

We have therefore $(I - (A_1 + A_2)\Delta t)\, c^{n+1} = c^n$. The method is based on approximations such as $\big(I - (A_1 + A_2)\Delta t\big) \simeq (I - A_2\Delta t)(I - A_1\Delta t) + O(\Delta t^2)$.

3.2 Time integration of chemical kinetics

In this subsection, Δt is the subcycling timestep.

The CTMs are usually based on the so-called Method of Lines: the space discretization is first performed (especially with finite differences or finite volumes methods) and then the time integration of the resulting ODE. Let us consider the case of reactive dispersion (the processes are advection, diffusion and chemistry). The time integration of the advection step is not an issue (or specified by the advection scheme); the time integration of (linear) diffusion is usually performed with an implicit solver; the most challenging point is then the time integration of chemical kinetics.

Let the system of ODE describing chemical kinetics be:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = \chi_i(c, t)\,, \quad 1 \leq i \leq n \tag{11}$$

The first difficulty is the large dimension ($n \gg 1$): the dimension of a comprehensive gas-phase chemical mechanism is typically of magnitude 100 (for instance, RACM [133]). The second difficulty is related to the wide range of timescales, which induces the so-called *stiffness* [15]. Benchmarks of explicit and implicit methods have been performed in many works. One refers for instance to [112–114, 148].

Among many candidates, the second-order Rosenbrock solver ([151] for gas-phase chemistry; [35] for aqueous-phase chemistry) is a promising algorithm, as it realizes a good trade-off between a reasonable accuracy and a rather low computational burden. The Rosenbrock method is defined in the following way from iteration n to iteration $n+1$:

$$\begin{cases} c^{n+1} = c^n + \dfrac{k_1 + k_2}{2} \\ (I - \gamma \Delta t\, J)k_1 = \Delta t \chi(c^n) \\ (I - \gamma \Delta t\, J)k_2 = \Delta t \chi(c^n + \Delta t\, k_1) - 2\gamma \Delta t\, J\, k_1 \end{cases} \tag{12}$$

$J = \frac{\partial \chi}{\partial c}$ is the Jacobian matrix for chemical kinetics. $\gamma = 1 + 1/\sqrt{2}$ ensures the L-stability of the method (see [149]). Preprocessing the computation of J with appropriate techniques that take into account the sparse nature (a chemical species is actually involved in a small set of reactions) is a powerful technique [24, 77, 110].

Another candidates are provided by explicit and quasi-explicit methods. The so-called quasi steady-state assumption (QSSA) methods ([165] for the original paper and then, for instance, [64, 86, 153]) are still widely used even if they do not perform well in fair benchmarks (with error/CPU diagrams). Such algorithms are strongly related to the so-called production-loss form of (11):

$$\chi_i(c) = P_i(c) - L_i(c)c_i \tag{13}$$

with P_i and L_i two nonnegative terms. This form is easy to get by partitioning the chemical reactions into two sets: reactions in which the species i is a reactant (loss term) and reactions in which it is a product (production term). The fast species and the slow species are defined by timescales $L_i \gg 1$ and $L_i \ll 1$, respectively. There are many variations on the basis of this partitioning of species. In the simplest form of the QSSA method, the fast species are computed from iteration n to iteration $n+1$ with the equilibrium $c_i^{n+1} = \frac{P_i}{L_i}$, whereas the slow species are computed with $c_i = \frac{P_i}{L_i} + (c_i^n - \frac{P_i}{L_i}) \exp(-L_i \Delta t)$.

Beyond QSSA methods, quasi-explicit methods are also very popular. The two-step method [146, 150] is an interesting algorithm that can be seen as an extension of the Euler backward iterative method (EBI) [57, 97]. Using the backward Euler method leads to the following nonlinear algebraic system:

$$\frac{c_i^{n+1} - c_i^n}{\Delta t} = \chi_i(c^{n+1}) \tag{14}$$

This can be solved by an iterative algorithm, for instance, the Newton algorithm, which requires to compute the Jacobian matrix J. Another approach is to approximate $\chi_i(c^{n+1}) \simeq P_i(c^n) - L_i(c^n)c_i^{n+1}$. The algebraic system can then be solved with:

$$c_i^{n+1} = \frac{c_i^n + P_i(c^n)\Delta t}{1 + L_i(c^n)\Delta t} \tag{15}$$

Variations of this idea lead to the EBI (first-order) and Twostep (second-order) algorithms. Notice that $P(.)$ and $L(.)$ can be evaluated with the components of c_i^{n+1} that have been already computed (which is the starting point of many tailored techniques).

Let us mention for the sake of completeness the combined explicit/implicit methods (for instance, [50]), which are based on the partitioning of chemical species in a set of slow species (integrated with an explicit solver) and a set of fast species (integrated with an implicit solver). We will go back to this approach in Section 7.

To conclude, a specific issue for the time integration of chemical kinetics is related to the positivity of concentrations. A usual tricky approach is the *clipping* (set the negative numerical concentrations to 0 or to a small predefined parameter). We refer to [107] for a rigorous method.

3.3 Some points related to advection

Solving advection is the other challenging step for the CTMs. The results strongly depend on the accuracy of the wind velocity (computed in meso-scale meteorological models). There are, however, some remaining issues for CTMs.

The first point is that the advection has to take care of sharp gradients. This is a key concern for the dispersion of accidental releases, for instance, of radionuclides. There is indeed no background field, and the point accidental emission implies strong gradients. Advection solvers with flux limiting are therefore often advocated (see a general presentation in [149]).

Another specific point is related to the so-called mass consistency problem [131] related to the off-line coupling between atmospheric dispersion models and meso-scale numerical models that provide meteorological fields such as wind velocity.

The continuity equation for density ρ can be written under the form $\dfrac{\partial \rho}{\partial t} + \mathrm{div}(V\rho) = 0$ with V the wind field.

The advection of a passive tracer of concentration c (defined, for instance, as the ratio of mass per air volume) is given by $\frac{\partial c}{\partial t} + \mathrm{div}(Vc) = 0$. The evolution of the mass mixing ratio $q = c/\rho$ (mass of species per air mass) is then given by $\frac{\partial q}{\partial t} + V \cdot \nabla q = 0$.

The conservation of uniform mixing ratio in the case of passive tracers is then easy to prove: if $q(0) = 1$, one gets for any time t, $q(t) = 1$. This property is usually referred as *mass consistency*. The key point is the consistency between the wind field used for ρ and the wind field used for c. The discrepancy may result from different numerical algorithms and discretization strategies between the computation of the wind velocity (inside the meteorological model) and the advection step (inside the CTM).

We briefly present three possible methods to circumvent these difficulties.

A first approach is the renormalization technique. Let $c^{n+1} = \mathcal{S}(c^n, V^n)$ be the advection scheme used by the CTM (supposed to be an explicit solver, which is usually the case in practice). Let us define $\tilde{\rho}^{n+1} = \mathcal{S}(\rho^n, V^n)$ the density that is computed with this scheme (this is not the "true" density ρ^{n+1} if the meso-scale model does not use the same scheme). The renormalization approach is then based on the correction:

$$c^{n+1} = \frac{\rho^{n+1}}{\tilde{\rho}^{n+1}} \mathcal{S}(c^n, V^n) \tag{16}$$

It is of course easy to check that mass consistency is met. One refers, for instance, to [16, 73]. A major drawback of this approach is that it may artificially introduce mass.

A second approach is based on the use of mass mixing ratio in the computation of fluxes after having noticed that $\frac{\partial c}{\partial t} + \mathrm{div}(\rho V \frac{c}{\rho}) = 0$. Convergence results for this algorithm are, however, not proved. One refer to [92] for the special case of the upwind scheme (without flux limiting).

A third approach is to modify the vertical component of wind velocity w to satisfy mass consistency. Suppose that horizontal advection is defined by $c_{\mathrm{H}} = \mathcal{S}_{\mathrm{H}}(c^n, u^n, v^n)$ (eventually with directional splitting). One defines $\rho_{\mathrm{H}} = \mathcal{S}_{\mathrm{H}}(\rho^n, u^n, v^n)$. Let the vertical advection scheme be given by $c^{n+1} = \mathcal{S}_V(c_{\mathrm{H}}, w^n)$. We therefore try to compute w^n such that the constraint $\rho^{n+1} = \mathcal{S}_V(\rho_{\mathrm{H}}, w^n)$ is met. Of course, this implies that this constraint may be solved in w^n, which implies that a nonlinear scheme (for instance, with flux limiting) cannot be used for vertical advection. One example of a scheme easy to "invert" is the upwind scheme [60, 89].

3.4 Grid resolution

In the case of point sources, having a fine resolution may be a key issue. We have already briefly mentioned the plume-in-grid approach. An alternative to this approach is of course to use fine meshes. A first method is based on adaptive gridding [132]. Unstructured meshes are more appropriate for this aim (see for instance [140]). In [143], local refinement is applied with a uniform-grid-based cascade technique that produces nested grids within one timestep. A second method is based on *nesting* techniques. The idea is to compute solutions on different fixed meshes of variable resolutions with one-way or two-way fluxes of informations (for instance, see [45]).

The grid resolution may have a strong impact on the results. The future generation of CTM will have a fine resolution (let say 0.1° for the horizontal dimension, less than 10 km) in coherence with the high-resolution meteorological models, currently under development.

4 Aerosol modeling

4.1 The General Dynamics Equation for aerosols

The numerical simulation of gaseous species is the first component of a CTM. For many applications, the modeling of condensed matter is also required. This includes the description of chemical composition and of size distribution of aerosols (particulate matter). Both have an influence for the radiative behavior of aerosols, for microphysical processes (through the activation of the finest aerosols to generate cloud droplets) and for the assessment of health impact (the most adverse particles are the finest ones).

Simulating a polydisperse distribution of aerosols is a challenging task for many reasons (at the numerical level):

- There is a wide range of scales (from a few nanometers to at least 10 μm).
- One has to deal with high-dimensional models (the number of chemical species inside the aerosol mixture has to be multiplied by the number of degrees of freedom chosen to describe the size).
- The processes that affect the size and chemical composition of aerosols are nonlinear and discontinuous (thresholds for phase transition).

The evolution of size distribution and of chemical composition for aerosols is governed by the so-called General Dynamics Equation for aerosols (GDE). It describes the impact of some processes (Fig. 2)

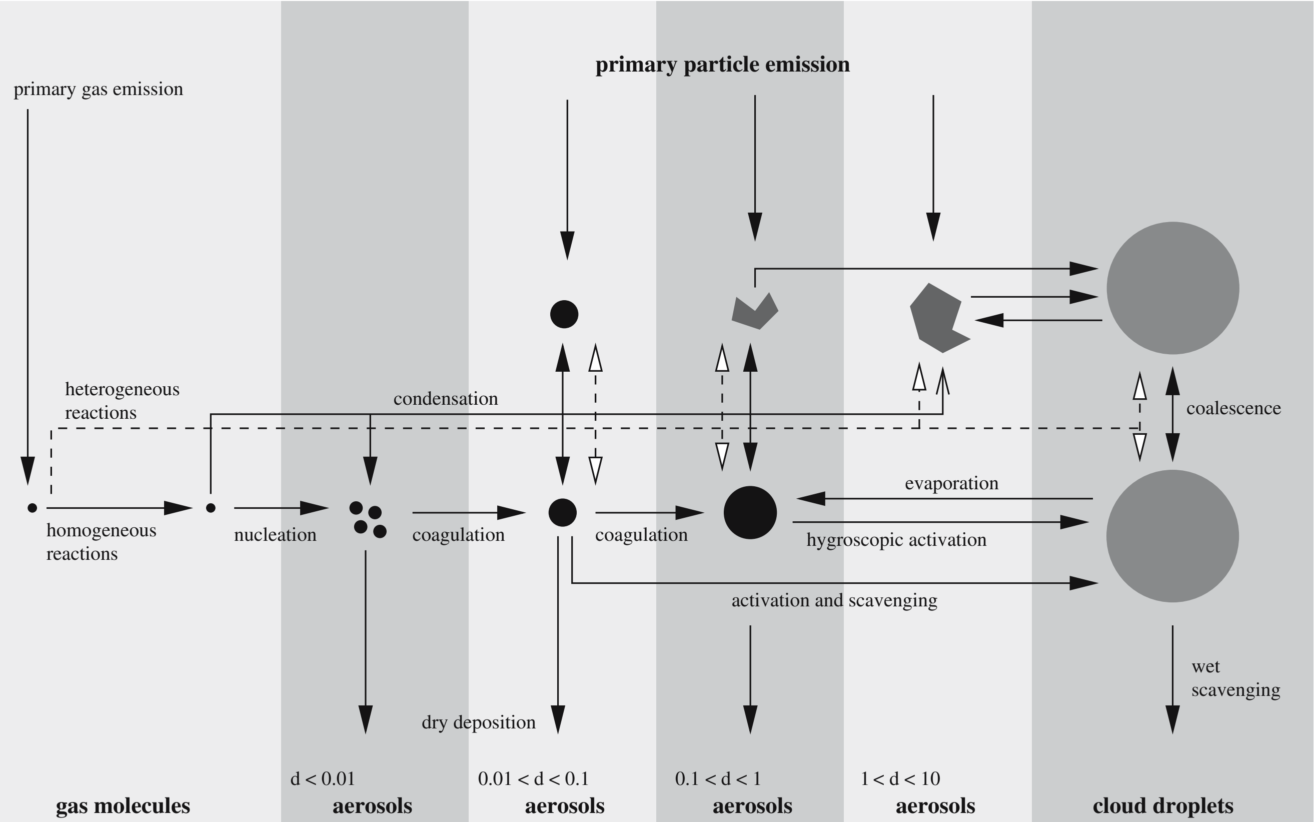

Fig. 2 Aerosol dynamics (the particle diameter *d* is in micrometers)

such as coagulation (a particle is created after a "collision" among two preexisting particles), condensation/evaporation (mass transfer between the aerosol phase and the gas phase) and nucleation (the formation of the smallest particles from stable clusters of aggregated gaseous molecules above a threshold, let say m_0 for mass).

Let $n(m,t)$ be the number distribution (for instance, with respect to aerosol mass) and $q_i(m,t)$ the mass distribution for species i (we use the same notation as for the mass mixing ratio in the previous section). The GDE is composed of the following equations:

- For the number distribution:

$$\frac{\partial n}{\partial t}(m,t) = \underbrace{\frac{1}{2}\int_{m_0}^{m-m_0} K(u,m-u)n(u,t)n(m-u,t)\mathrm{d}u}_{\text{coagulation gain}} - \underbrace{n(m,t)\int_{m_0}^{\infty} K(m,u)n(u,t)\,\mathrm{d}u}_{\text{coagulation loss}} - \underbrace{\frac{\partial(I_0 n)}{\partial m}}_{\text{mass transfer}} + \underbrace{\delta(m_0,m)J_0(t)}_{\text{nucleation}} \qquad (17)$$

- For the mass distributions:

$$\frac{\partial q_i}{\partial t} = \int_{m_0}^{m-m_0} K(u,m-u)q_i(u,t)n(m-u,t)\,\mathrm{d}u - q_i(m,t)\int_{m_0}^{\infty} K(m,u)n(u,t)\,\mathrm{d}u - \frac{\partial(I_0 q_i)}{\partial m} + I_i(m_1,\ldots,m_{n_c},t)n(m,t) + \underbrace{\chi_i^{aer}(m_1,\ldots,m_{n_c})n(m,t)}_{\text{Internal chemistry}} \qquad (18)$$

An aerosol of mass m is composed by n_c components of mass $m_i(m)$, such that $\sum_{i=1}^{i=n_c} m_i(m) = m$. Notice that $q_i(m) = m_i(m)n(m)$. An aerosol with a given mass is assumed to have a unique composition (*internal* mixing).

$K(.,.)$ is the coagulation kernel (due to Brownian motion, not detailed here). I_i is the condensation/evaporation rate for volatile species i:

$$I_i = a_i(m)\left(c_i - \eta_i\, c_i^{eq}(m_1(m),\ldots,m_{n_c}(m))\right) \qquad (19)$$

where $a_i(m)$ stands for a parameterization of a mass transfer coefficient (a function of accommodation, of molecular diffusion and of particle size), c_i is the gas-phase concentration of species i and c_i^{eq} is the surface concentration of species i (supposed to be at equilibrium with the aerosol mixture). η_i is the Kelvin coefficient that describes the curvature effect. $I_0 = \sum_i I_i$ is the total rate.

J_0 is the nucleation rate, given by parameterizations as a function of thermodynamic conditions (temperature and relative humidity) and of some gaseous concentrations (water, ammonium, sulfate).

The internal chemistry is usually neglected.

This set of partial integro-differential equations is strongly coupled, especially through the condensation/evaporation term. One key component is the availability of a thermodynamic model for computing c_i^{eq} (it is the most expensive part of the model, as far as CPU is concerned) [169].

Notice that the GDE is usually not solved simultaneously with the other processes. We refer to [84] for the investigation of splitting errors related to microphysical-multiphase chemical systems.

4.2 Stochastic methods

Analytical solutions of the GDE only exist for some academic cases [48, 102]. This means that it is difficult to benchmark numerical algorithms with respect to a reference (exact) solution. The lack of reference solution has already been underlined in [142]. Searching for a numerical reference solution is therefore a key issue. Such a solution is not necessarily computationally efficient but has to be highly accurate.

The stochastic methods are good candidates for being such reference solutions. They have a rigorous theoretical basis (through the stochastic formulation of the GDE) and are easy to implement. In stochastic methods, one deals with *numerical particles* that represent, in a way to specify (see below), the aerosol distribution.

We refer to [37] for the application to the coagulation equation (the so-called Smoluchowski equation). A numerical particle is associated to a fixed number of physical aerosols of a given size. The drawback is that the number of numerical particles decreases as coagulation makes the number of aerosol decrease. This is a strong limitation as the method converges in the limit of the number of particles going to infinity. [1, 38] have associated to one numerical particle a fixed mass of aerosols. As coagulation preserves mass, the number of numerical particles can remain constant. This kind of algorithms is usually referred as *mass flow algorithms*.

We refer to [30] for the extension to the GDE. The key idea is to use mass flow algorithms with a varying mass due to condensation/evaporation.

4.3 Modal methods

For 3D modeling, a popular approach is based on the *modal* approach. The aerosol distribution is represented by a given number of modes (typically log-normal functions of aerosol mass or volume) [159]. For instance in the Modal Aerosol Model (MAM, [115]), one usually defines three modes in the fine range (nucleation i, Aitken j and accumulation k) and one mode in the coarse range (c). Each of these modes corresponds to a distinct range of particle dry diameters D: $D < 0.01\ \mu m$ for the nucleation mode, $0.01\ \mu m < D < 0.1\ \mu m$ for the Aitken mode, $0.1\ \mu m < D < 2.5\ \mu m$ for the accumulation mode and $D > 2.5\ \mu m$ for the coarse mode. Each of the four modes has its own chemical composition.

The aerosol distribution is then given by summing the mode contributions, for example, for the number distribution: $n(D,t) = n_i(D,t) + n_j(D,t) + n_k(D,t) + n_c(D,t)$, where n_i is the log-normal distribution for mode i, etc. For $l = i, j, k, c$, one has:

$$n_l(D) = \frac{N_l}{\sqrt{2\pi}\,\ln(\sigma_l)\ d_l}\exp\left[-\frac{1}{2}\left(\frac{\ln^2(D/d_l)}{\ln^2(\sigma_l)}\right)\right] \tag{20}$$

where N_l is the total number of particles within mode l, d_l is the "dry" median diameter and σ_l is the standard deviation. The moment of order h of the distribution is defined as

$$M_h^l = \int_{-\infty}^{\infty} D^h\, n_l(D)\ \mathrm{d}(D) \ = \ N_l\, d_l^h\ \exp\left(\frac{h^2}{2}\ln^2\sigma_l\right). \tag{21}$$

The modal distribution is known once the three parameters N, d_g and σ_g are known. For each mode, an evolution equation is derived from the GDE for three moments (M_0, M_3 and M_6), from which the three parameters may be computed:

$$N = M_0,\ d_g = \left(\frac{M_3^4}{M_6 M_0^3}\right)^{\frac{1}{6}},\ \sigma_g = \exp\left(\sqrt{\frac{1}{9}\ln\left(\frac{M_0 M_6}{M_3^2}\right)}\ \right) \tag{22}$$

The third moments are computed for each chemical species to follow the chemical composition in each mode.

The resulting system of (stiff) ODEs can be solved with appropriate time integration solvers (see previous section).

A specific issue for the 3D implementation of modal methods is related to the management of mode *merging* and *splitting*. Merging between two modes is required to force modes to be of distinct size ranges throughout the simulations [9, 160]. When the combined effect of nucleation and condensation and the effect of coagulation are of the same order of magnitude but act in opposite direction, a mode may split into two modes. This is particularly likely to happen for the nucleation mode i during high nucleation episodes. A splitting scheme has then to be applied [9].

4.4 Size-resolved methods

Modal methods are based on an a priori analytical form of the aerosol distribution. Another approach is of course to discretize the GDE, viewed as a system of integro-differential equations. We briefly present variational methods (based on a weak formulation of the GDE) and the more classical sectional method (based on finite differences).

4.4.1 Variational methods

Let us write the GDE under the following form:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = f(n, x, t)\ , \quad \frac{\mathrm{d}q_i}{\mathrm{d}t} = f_i(n, q_i, x, t) \tag{23}$$

$x = \ln m$ is the logarithm of the dry mass (the variable that is classically used for numerical algorithms). If V is a functional space (of functions of x) with an appropriate scalar product $(.,.)$, one defines the weak formulation of the GDE (for instance, for the number distribution) as:

$$\forall v \in V,\ \left(\frac{\mathrm{d}n}{\mathrm{d}t}, v\right) = (f(n, x, t), v) \tag{24}$$

A numerical method is then based on a projection onto a subspace V_k of V, of finite dimension k:

$$\Pi n(x, t) = \sum_{j=1}^{j=k} n^j(t) L_j(x) \tag{25}$$

where the functions $L_j(x)$ are the basis functions of V_k.

In [95], the basis functions are the so-called “hat functions” and one gets a finite element method. Another family of functions with a bounded domain of definition is given by cubic splines [47, 106].

When the basis functions do not have a bounded domain of definition, for instance, in $[x_0, \infty[$, one usually refers the method as a spectral method. In [109], the basis functions are polynomial functions of Lagrange. In [25], they are polynomial functions of Laguerre and of Legendre.

Apart from the choice of the basis functions L_j, the test functions v have also to be chosen. In the case of a *collocation* method, the test functions are Dirac functions at collocation points $(x_i)_{i=1,\ldots,k}$: $v_i(x) = \delta(x - x_i)$. For a basis of Lagrange polynomials, one easily gets for the number distribution [27, 109]:

$$\frac{\mathrm{d}\vec{n}}{\mathrm{d}t} = \underbrace{A(\vec{n})\vec{n} - \mathrm{diag}(B\vec{n})\vec{n}}_{\text{coagulation}} - \underbrace{C\vec{n}}_{\text{c/e}} \tag{26}$$

where $(\vec{n})_j = n^j$. The matrices $A(n)$ (the line i of $A(n)$ is under the form $\vec{n}^{\mathrm{T}} A_i \vec{n}$ with A_i a matrix) and B are related to coagulation. The matrix C is related to condensation/evaporation.

The resulting system of ODEs has then to be integrated with appropriate solvers. We refer to [109] and to [27] for the benchmark of two collocation methods.

The advantage of such methods is that one can solve the GDE without using splitting methods. However, such methods may suffer from numerical instabilities and are computationally demanding.

4.4.2 Sectional methods

Another approach is based on finite differences. The aerosol distribution is discretized in a fixed number of *bins* (let say n_{b}) or *sections*. This justifies the name of *sectional* method. The bins correspond to the cells in the classical terminology of numerical analysis.

The objective is then to compute Q_i^j and N^j, the integrated values over the bin j for the mass distribution of species i and the number distribution, respectively.

Splitting of coagulation and condensation/evaporation is often used. For the sake of clarity, we use splitting in our presentation.

4.4.3 Coagulation

Here, we only present the equations for the number distribution. After integration of the coagulation equation over bins, one typically gets:

$$\frac{\mathrm{d}N^j}{\mathrm{d}t}(t) = \frac{1}{2}\sum_{j_1=1}^{j}\sum_{j_2=1}^{j} X^j_{j_1 j_2} N^{j_1} N^{j_2} - N^j \sum_{j_1=1}^{n_{\mathrm{b}}} X_{j j_1} N^{j_1} \tag{27}$$

The coefficients $X^j_{j_1 j_2}$ are the key components of the method and describe the fraction of the aerosol generated by the coagulation of bins j_1 and j_2 located in bin j. $X_{j_1 j_2}$ is defined by summing over all possible resulting bins: $X_{j_1 j_2} = \sum_{j=1}^{j=n_b} X^j_{j_1 j_2}$.

A popular method is based on [63, 135]. We refer to [29] for a review and a rigorous derivation of such equations. The key point is to choose the closure scheme inside each bin (the a priori local form of the aerosol distribution).

The resulting equation is then integrated with an ODE solver. A usual approach is based on a semi-implicit approach, for instance:

$$(N^j)^{n+1} = \frac{(N^j)^n + \frac{\Delta t}{2}\sum_{j_1=1}^{j-1}\sum_{j_2=1}^{j} X^j_{j_1 j_2}(N^{j_1})^{n+1}(N^{j_2})^n}{1 + \Delta t\sum_{j_1=1}^{n_b}(X_{jj_1} - X^j_{jj_1})(N^{j_1})^n} \tag{28}$$

4.4.4 Condensation/evaporation

If the bins are fixed, the method is referred as an *Eulerian* method. A crucial point is to lower the numerical diffusion (condensation/evaporation may be viewed as an advection process). Classical methods are used in this aim (for instance, semi-Lagrangian solvers) [23, 31].

An alternative approach is to use *Lagrangian* methods by following the evolution of bins according to condensation/evaporation. This method is also referred as a *moving sectional* algorithm [62, 68]. A motivation is that one can only use a small number of bins in 3D applications. A crucial point is then to map the resulting moving distribution onto a fixed grid after having solved the GDE because the aerosol distribution is given on a fixed grid for 3D processes, such as advection or diffusion.

From (19), one gets a system of ODEs that gives the time evolution of Lagrangian variables. For instance, for mass distributions:

$$\frac{\mathrm{d}Q_i^j}{\mathrm{d}t} = N^j a_i(m_j)\left(K_i - \sum_{j'=1}^{n_b} Q_i^{j'} - \eta_i^j\, c_i^{\mathrm{eq}}(Q_1^j, \ldots, Q_{n_c}^j)\right) \tag{29}$$

The mass conservation law for the total mass (gas and aerosol) is written under the form $c_i + \sum_{j'=1}^{n_b} Q_i^{j'} = K_i$.

This system is highly stiff because of the wide range of timescales for condensation/evaporation. We refer to the previous section for the issue of time integration. In the size resolved aerosol model (SIREAM, [26]), a Rosenbrock method is used. Some specific flux limitations can also be applied to reduce the stiffness (especially this because of the fluxes of H^+) [96].

We refer to [168] for a comparative review of algorithms.

5 Inverse modeling and data assimilation

5.1 Background

We now assume that the model has been defined (that is to say that parameterizations have been proposed; see Section 2) and that appropriate numerical solvers are used (Sections 3 and 4).

A field of growing interest is the use of data assimilation techniques. The objective is to couple model outputs and observational data provided by monitoring networks to improve the forecasts (data assimilation) and/or to reduce the uncertainties of some input parameters (inverse modeling, for instance, of emissions). We refer, for instance, to [3, 66, 137] for a general presentation in geophysics.

Let use write the evolution system (1) under the form:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = f(c, \Psi)\,, \quad c(0) = c_0 \tag{30}$$

Here, Ψ stands for some uncertain forcing parameters (for instance, emissions or boundary conditions). The uncertainties are related to f (through the parameterizations), to the initial conditions c_0 and to the forcing parameters Ψ. Notice that in meteorology, the key point is the sensitivity with respect to initial conditions, because of the chaotic behavior of atmospheric dynamics (for instance, [66]). In atmospheric chemistry, the sensitivity with respect to the initial conditions is much lower because of the dissipative nature of chemical kinetics. A typical spin-up is of magnitude 1 week at continental scale. Notice that in atmospheric chemistry, the eigenvalues of $J = \frac{\partial f}{\partial c}$ have strictly negative real parts. This is not a wave problem as in meteorology.

Some observational data may be provided by ground stations (for instance, the local monitoring networks) or satellital data related to earth observation. Let us assume that some observations $\mathrm{obs}^1, \ldots, \mathrm{obs}^{n_o}$ are available at discrete times $t_1, \ldots, t_{n_o}$. At time t_n, the model output is $c(t_n) = F^n(\Psi)$ (we use this notation to underline the input/output relation). The observation is not necessarily the model output, and we write H^n the observation operator at time t_n: if there were no errors, one should get $\mathrm{obs}^n = H^n\, F^n(\Psi)$.

Data assimilation techniques are based on the reduction of discrepancies between the model outputs and the observational data, measured by the cost function of the model parameters Ψ (when required, it could also depend on the initial conditions):

$$J_{\mathrm{o}}(\Psi) = \sum_{n=1}^{n=n_{\mathrm{o}}} \underbrace{(\mathrm{obs}^n - H^n\, F^n(\Psi))^T R_n^{-1} (\mathrm{obs}^n - H^n\, F^n(\Psi))}_{J_n} \tag{31}$$

R_n is the covariance matrix for observation error (to be specified). A second part of the cost function can be defined by taking into account the so-called background term (an a priori knowledge of Ψ):

$$J_{\mathrm{b}}(\Psi) = (\Psi - \Psi_{\mathrm{b}})^T B^{-1} (\Psi - \Psi_{\mathrm{b}}) \tag{32}$$

with Ψ_{b} the background value and B the background error covariance matrix. Adding J_{b} to J_{o} is a penalization technique (Tikonov regularization) that may be necessary for ill-posed (high-dimensional) problems.

The problem is then to minimize $J = J_{\mathrm{o}} + J_{\mathrm{b}}$ with respect to $\Psi \in V_{\mathrm{adm}}$. This provides an optimized value $\Psi^\star = \mathrm{argmin} J$ that can be used to compute, when required, an optimal forecast $c^\star = F(\Psi^\star)$. V_{adm} is the space of admissible parameters (for instance, some parameters have nonnegative values).

There are many applications of this approach in inverse modeling of emissions: for CH_4 (for instance [5]), for carbon ([21] at regional scale and [4] at global scale), for NO_x [99] and VOC [20]. Another related problem is the so-called localization problem for point sources of accidental releases (for instance, radionuclides: [10, 11]).

5.2 Sequential versus variational methods

Two families of methods are widely used: the variational algorithms, based on a minimization procedure (3D-Var when time evolution is not taken into account, and 4D-Var otherwise), and the sequential methods, based on the Kalman filter theory.

5.2.1 Variational methods

The variational methods are based on gradient-like minimization methods (for instance, [111]). They require to compute $\nabla_\Psi J$. The finite differences techniques are not useable because of the computational burden, and one often advocates the use of the so-called *adjoint models*. The CPU cost required for the adjoint model ranges typically from three to seven times the CPU cost required for the model.

Let us detail the algorithm for a CTM. Here, we assume that the CTM (actually the computation of the cost function) is given by the sequence:

1. Initialization of the cost function: $J = 0$ (or a background term)
2. Time loop labelled by $1 \leq n \leq n_{\mathrm{o}}$:
    - Get the forcing fields ϕ^{n-1}
    - Compute the new state $c^n = F(c^{n-1}, \phi^{n-1})$
    - Update the cost function: $J = J + J_n(c^n)$

The forcing fields ϕ^n include the meteorological data, the results of parameterizations, etc.

A direct calculation of ∇J could be performed, but the storage requirement would be too high because $(\phi^0, \ldots, \phi^{n_{\mathrm{o}}-1})$ and the trajectory $(c^0, \ldots, c^{n_{\mathrm{o}}})$ should be stored. The idea is to notice that $\nabla_\Psi J = \sum_n \nabla_\Psi J_n$, and:

$$\nabla_\Psi J = \sum_{1 \leq m < n \leq n_{\mathrm{o}}} \left(\frac{\partial c^m}{\partial \Psi}\right)^{\mathrm{T}}_{|t_{m-1}} \cdots \left(\frac{\partial c^n}{\partial c^{n-1}}\right)^{\mathrm{T}}_{|t_{n-1}} \left(\frac{\partial J_n}{\partial c^n}\right)^{\mathrm{T}}_{|t_n} + \sum_{1 \leq n \leq n_{\mathrm{o}}} \left(\frac{\partial c^n}{\partial \Psi}\right)^{\mathrm{T}}_{|t_{n-1}} \left(\frac{\partial J_n}{\partial c^n}\right)^{\mathrm{T}}_{|t_n} \tag{33}$$

The adjoint model gives, for any vector v, $(\frac{\partial c^n}{\partial \Psi})^{\mathrm{T}}_{|t_{n-1}} v$ and $(\frac{\partial c^n}{\partial c^{n-1}})^{\mathrm{T}}_{|t_{n-1}} v$. If the differentiation is performed with respect to the initial conditions ($\Psi = c_0$), the formula is:

$$\begin{aligned}\nabla_{c_0} J &= \sum_{1 \leq n \leq n_o} \left(\frac{\partial c^1}{\partial c^0}\right)^T_{|t_0} \cdots \left(\frac{\partial c^n}{\partial c^{n-1}}\right)^T_{|t_{n-1}} \left(\frac{\partial J_n}{\partial c^n}\right)^T_{|t_n} \\ &= \ldots \times \left[\left(\frac{\partial J_{n_o-2}}{\partial c^{n_o-2}}\right)^T + \left(\frac{\partial c^{n_o-1}}{\partial c^{n_o-2}}\right)^T \left[\left(\frac{\partial J_{n_o-1}}{\partial c^{n_o-1}}\right)^T \right.\right. \\ &\qquad \left.\left. + \left(\frac{\partial c^{n_o}}{\partial c^{n_o-1}}\right)^T \left(\frac{\partial J_{n_o}}{\partial c^{n_o}}\right)^T\right]\right]\end{aligned} \tag{34}$$

In this case, the algorithm for computing $\nabla_{c_0} J$ is then:

1. Initialization: $(c^{n_{\mathrm{o}}})^\star = \left(\frac{\partial J_{n_{\mathrm{o}}}}{\partial c^{n_{\mathrm{o}}}}\right)^{\mathrm{T}}$,
2. Reverse time loop labelled by $n_{\mathrm{o}} \geq n \geq 1$:
    - Get the forcing fields ϕ^{n-1}
    - Get the stored trajectory c^{n-1}
    - Update $(c^{n-1})^\star = \left(\frac{\partial c^n}{\partial c^{n-1}}\right)^{\mathrm{T}}_{|t_{n-1}} (c^n)^\star + \left(\frac{\partial J_{n-1}}{\partial c^{n-1}}\right)^{\mathrm{T}}$
3. Conclude with: $\nabla_{c_0} J = (c^0)^\star$

A specific issue is related to the availability of the adjoint model. There are two solutions: the adjoint model may be computed on the basis of the discretization of the continuous adjoint model, or it may be given by the adjoint code of the discretized model. We refer for instance to [120].

In many cases, automatic differentiation with appropriate tools is performed [49, 52, 77]. Notice that discontinuities exist in the CTM (especially related to aerosol thermodynamics with thresholds) and may require a special care.

We refer for instance to [39–42, 99] for the application of variational methods to air quality modeling at regional scale.

In the linear case, when the chemical production term is given by a linear term ($\chi(c) = \chi\ c$, with χ a matrix), a simplification can be used. For the sake of clarity, we consider the continuous case. The cost function is then given over an assimilation window $[0, T]$ as $J(c) = \int_0^T j(c(t))$. For instance, we suppose that the control parameter is the source term S. Let us define the adjoint variable $c^\star$ governed by (we omit density for the sake of clarity):

$$-\frac{\partial c_i^\star}{\partial t} - V\mathrm{div}(c_i^\star) = \mathrm{div}(K\nabla c_i^\star) + (\chi^{\mathrm{T}} c^\star)_i - \Lambda_i {c^\star}_i + \nabla_c j\,,\quad c_i^\star(T) = 0 \tag{35}$$

When the wind field is divergence-free, one gets:

$$-\frac{\partial c_i^\star}{\partial t} - \mathrm{div}(V c_i^\star) = \mathrm{div}(K\nabla c_i^\star) + (\chi^{\mathrm{T}} c^\star)_i - \Lambda_i {c^\star}_i + \nabla_c j\,,\quad c_i^\star(T) = 0 \tag{36}$$

The key result is that $\nabla_{S(t)} J = c^\star(t)$. The initial model can be used to compute the sensitivity by modifying the input parameters and performing a backward integration (compare (1) and (36)). The so-called *retroplume* methods (for instance, [59]) are based on this approach.

5.2.2 Sequential methods

The sequential methods are based on the Kalman filter theory (for instance, [137]). We distinguish the forecast and the analysis obtained by taking into account the observational data. Both are supposed to be stochastic variables that follow normal laws, $\mathcal{N}(c_{\mathrm{f}}, P_{\mathrm{f}})$ and $\mathcal{N}(c_{\mathrm{a}}, P_{\mathrm{a}})$, respectively. P is a covariance error matrix. The model is supposed to be given by $c^{n+1} = F(c^n)$ from t_n to t_{n+1}, with an unbiased model error following $\mathcal{N}(0, Q)$.

The most basic algorithm (BLUE for Best Linear Unbiased Estimator and the extension to the nonlinear case with the extended Kalman filter) is given by the following sequence:

1. Forecast from t_n to t_{n+1}:

$$c_{\mathrm{f}}^{n+1} = F(c_{\mathrm{a}}^n) \tag{37}$$

2. Propagation of the error covariance matrix:

$$P_{\mathrm{f}}^{n+1} = M_{n,n+1} P_{\mathrm{a}}^n M_{n,n+1}^{\mathrm{T}} + Q^n \tag{38}$$

3. Analysis at t_{n+1}:

$$c_{\mathrm{a}}^{n+1} = c_{\mathrm{f}}^{n+1} + K^{n+1}\left(\mathrm{obs}^{n+1} - H^{n+1} c_{\mathrm{f}}^{n+1}\right) \tag{39}$$

where the gain matrix is:

$$K^{n+1} = P_{\mathrm{f}}^{n+1}(H^{n+1})^{\mathrm{T}}\left(H^{n+1} P_{\mathrm{f}}^{n+1}(H^{n+1})^{\mathrm{T}} + R_{n+1}\right)^{-1} \tag{40}$$

4. Computation of the analysis covariance matrix at t_{n+1}:

$$P_{\mathrm{a}}^{n+1} = \left(I - K^{n+1} H^{n+1}\right) P_{\mathrm{f}}^{n+1} \tag{41}$$

$M_{n,n+1}$ is the linear tangent model from t_n to t_{n+1}.

This method is easy to implement but a key issue is the large dimension for reactive problems and the computational burden associated to the step (38). A powerful technique is given by the use of *reduced* filters.

Notice that the covariance error matrices can be written as $P_{\mathrm{f}} = S_{\mathrm{f}} S_{\mathrm{f}}^{\mathrm{T}}$, $P_{\mathrm{a}} = S_{\mathrm{a}} S_{\mathrm{a}}^{\mathrm{T}}$, $Q = S_q S_q^{\mathrm{T}}$ and $R = S_{\mathrm{r}} S_{\mathrm{r}}^{\mathrm{T}}$. It is straightforward to check that $S_{\mathrm{f}} = (M S_{\mathrm{a}}\ ,\ S_q)$, $K = S_{\mathrm{f}}\ A(A^{\mathrm{T}}\ A + R)^{-1}$ and $S_{\mathrm{a}} = S_{\mathrm{f}}(I - A(R + A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}})^{1/2}$ with $A^{\mathrm{T}} = H S_f$. The advantage of using the square matrices S rather than the covariance matrices P is that these matrices are symmetric and strictly positive. Moreover, with $S = (s_1 \ldots s_{n_s})$ (n_s depends on the matrix, s_i is a column), one gets:

$$S S^{\mathrm{T}} = \sum_{i=1}^{n_s} s_i {s_i}^{\mathrm{T}} \tag{42}$$

Once the $s_i^{\mathrm{T}} s_i$ values have been ranked in a decreasing order, it is easy to define truncated matrices related to S.

For the RRSQRT filter [145], the modes are defined from the leading eigenvalues of $S_{\mathrm{a}}^{\mathrm{T}} S_{\mathrm{a}}$ (that has the same eigenvalues as P_{a}). For the SEEK (singular evolutive extended Kalman, [94]) and the SEIK (singular evolutive interpolated Kalman, [93]) filters, the truncation is based on the empirical orthogonal functions (EOF) of a computed trajectory.

An alternative approach is to use an *ensemble* filter by computing the covariance matrix on the basis of a Monte Carlo method:

$$P_f \simeq \frac{1}{N-1} \sum_{j=1}^{j=N} (c^{(j)} - \bar{c})(c^{(j)} - \bar{c})^{\mathrm{T}} \tag{43}$$

with an ensemble of N possible states $\left(c^{(j)}\right)$ of average $\bar{c}$.

We refer for instance to [116] for a benchmark of Kalman filters applied to air pollution modeling and also to [22] for the application of ensemble Kalman filters.

5.3 Coming issues

Let us conclude this section with two issues of growing interest.

5.3.1 Second-order sensitivity

The first issue is related to the so-called *second-order sensitivity*. Once the optimized parameter $\Psi^\star$ (for instance, emissions) has been computed, one can investigate its robustness with respect to other uncertainties. For instance, if Φ stands for uncertain parameters that are not optimized (because of the large dimension, for instance), $\Psi^\star$ is defined as a function of Φ:

$$\Psi^\star(\Phi) = \mathrm{argmin} J(\Psi, \Phi) \tag{44}$$

The second-order sensitivity [158] is defined as:

$$\frac{\partial \Psi}{\partial \Phi} = -\,(\mathrm{Hess\ J})^{-1} \frac{\partial \nabla_\Psi J}{\partial \Phi} \tag{45}$$

with the use of the implicit function theorem. We refer, for instance, to [100] with the application to a simple Gaussian model for the atmospheric dispersion of radionuclides.

Computing the Hessian matrix of J may be a difficult task for high-dimensional systems. A brute-force approach is then to investigate the sensitivity of $\Psi^\star$ with respect to modified model configurations or model parameters. An example is given by the inverse modeling of NO_x at regional scale and the sensitivity analysis with respect to parameterizations in [99]. A similar approach is also applied for the inverse modeling of CH_4 [83]. Notice that this point is crucial to use the optimized parameters for themselves and not only for forecast.

5.3.2 Network design

An interesting question is related to what is usually called *network design*: how to optimize a mobile or a fixed monitoring network (related to the observational operator H). This is a classical issue, for instance, in meteorology. Mathematically speaking, the quality of a monitoring network can be estimated by the condition number of the Hessian matrix of J (that depends on the network through H).

We refer, for instance, to [8, 130]. A preliminary application to air pollution modeling can also be found in [108].

6 Uncertainty propagation and ensemble forecast

6.1 Background

Air quality modeling and simulation suffer from many uncertainties (for instance, [105]):

- Many input data are poorly known (see, for instance, Table 1), especially emissions. Moreover, some forcing fields, such as meteorological fields, are computed with numerical models that may be uncertain.
- The models are highly parameterized (see Section 2).
- The numerical algorithms and the discretization also induce uncertainties (remember that the numerical resolution may be coarse because of the computational burden, especially for aerosols).

It is therefore not relevant to view outputs of CTMs as deterministic values. Even if the models are “validated” (it is more rigorous to say that model-to-data comparisons are performed, when possible), one must keep in mind that there are a large amount of degrees of freedom (especially in parameterizations) and only a small number of model outputs can be measured. For instance, most of the existing CTMs have been extensively tuned to meet acceptable model-to-data error statistics for ozone peaks at ground. It does not ensure that the results are satisfactory for 3D fields and many other trace species. There is therefore a danger of using “overtuned” models, especially for impact studies or long-term scenario studies.

6.2 Some possible strategies for assessing the impact of uncertainties

6.2.1 Sensitivity analysis

A first strategy is to evaluate the sensitivity of some model outputs, let say $c = F(\Psi)$, with respect to uncertain input parameters, let say Ψ, through the partial derivative $\frac{\partial F}{\partial \Psi}$.

Table 1 Uncertainties for some input parameters of CTM (parameters of log-normal probability density functions, following [54])

Variable	Range (95% of data)	σ (log-normal)
Ozone initial conditions	3	0.549
NO_x/VOC initial conditions	5	0.805
Major point NO_x/VOC emissions	1.5	0.203
Rainfall amount	2	0.347
Cloud liquid water content	2	0.347
Biogenic NO_x/VOC emissions	2	0.347

This approach has been widely used with many variations in the way the sensitivities are computed. We can cite among many other works: [18, 24, 55, 111, 117] for a general presentation, [32, 91, 166] for aqueous-phase chemistry, [78] for the sensitivity of ozone with respect to emissions, [167] for the sensitivity of ozone and [104] for the sensitivity of atmospheric mercury at European scale with respect to the boundary conditions, etc.

Notice that this item is also related to variational data assimilation (see Section 5) because the same techniques can be used to compute gradients and sensitivities.

This approach provides an estimation of the local uncertainty.

6.2.2 Monte Carlo simulations

A second approach is based on a classical Monte Carlo simulation on the basis of PDFs for the uncertain inputs.

The application to box models (for instance, chemical kinetics in [103]) is not too computationally demanding. In 3D applications, a major drawback is that the number of possible runs is usually low (typically 50 or 100). Because of the large dimension of systems, the brute-force approach is not feasible and some appropriate methods have to be used. A very attractive method is the so-called deterministic equivalent Monte Carlo method [138], which is based on polynomial expansions of the targeted outputs. For instance, for the simple case of one input ($\Psi \in \mathbf{R}$):

$$c = \alpha_0 + \sum_{i=1}^{i=N} \alpha_i \, P_i(\Psi) \tag{46}$$

where Ψ stands for the uncertain input parameter, P_i are orthogonal polynomial functions (“chaos” expansion) and N is the order of the approximation.

The coefficients (α_i) are computed at collocation points based on the PDF for Ψ (we refer to [138]). The essence of the algorithm is to simulate the model response surface by using orthogonal polynomials of the uncertain input parameters. Using (46) is much more efficient than using the exact model.

The application of this method to simulations over Europe for uncertainties in NO_x, NH_3 and SO_2 emissions can be found in [12, 14].

6.3 Multimodeling and ensemble forecast

As pointed out before, a key point is to avoid “overtuning” of models because of the large amount of output data and the small number of observed data. For instance, a typical version of a sectional aerosol model may include up to hundreds of species per grid cell (let say 10 bins times 10-20 chemical species per grid cell). In comparison, observational data are routinely available for aggregated values, such as PM2.5 or PM10 (PMx is the mass of aerosols whose aerodynamic diameter is less than or equal to x micrometers) but not for size distribution or chemical composition.

A promising approach is not to rely on one single model but to use a set (an *ensemble*) of models or of model configurations to deliver the forecast. We refer, for instance, to [80] for the study of model spreads due to uncertainties in model inputs, numerical algorithms and parameterizations.

The next step is to combine the results of the ensemble not only to assess the uncertainties but also to improve the forecast. We now detail a powerful approach based on what is usually called the *super-ensemble* technique [69].

Let $\mathcal{M}$ be the set of available 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). The simplest approach is to use the ensemble mean (EM):

$$\mathrm{EM}_{t,x} = \frac{1}{|\mathcal{E}|} \sum_{m \in \mathcal{E}} M_{m,t,x} \tag{47}$$

We refer, for instance, to [46] with the application to the forecast of radionuclides. Another similar approach is the use of the median model.

A promising approach is to take into account observations. Let $\mathrm{obs}_{t,x}$ be the observation corresponding to the model output $M_{t,x}$. An improved forecast may be given as a linear combination of model outputs as:

$$\mathrm{ELS}_{t,x} = \sum_m \alpha_m M_{m,t,x} \tag{48}$$

where α minimizes the cost function $\sum_{t,x}\left[\mathrm{obs}_{t,x} - \sum_m \alpha_m M_{m,t,x}\right]^2$, which measures a model-to-data discrepancy. The minimization can be performed in the least square sense (ELS stands for ensemble least square). The applications to air quality modeling can be found in [79, 90].

A challenging point is to forecast the weights α by only using *past* observational data. Some techniques based on machine learning and/or expert selection (some popular tools in statistics) can be used. One of the simplest methods uses the so-called loss function $L_t(\alpha_t) = \left(\sum_m \alpha_{m,t} M_{m,t} - \mathrm{obs}_t\right)^2$. The weights can be updated with a gradient algorithm [19] $\alpha_t = \alpha_{t-1} - \eta \nabla L_{t-1}(\alpha_{t-1})$. We refer, for instance, to [79].

7 Proxy and reduced models

7.1 Motivations

An important application of CTMs is the long-term estimation of the impacts of emissions reduction. This is, for instance, illustrated by the CAFE program in Europe (Clean Air For Europe). In such applications, the physical model is only a small part of a modeling chain that includes models for economic activities (and then emissions) and monetary evaluation of impacts due to air pollution. Such a modeling chain is used in an optimization mode to find the optimal partitioning of emissions that minimizes a global economic cost. This implies that the physical model (the CTM) has to be used many times.

It is of course not possible to use comprehensive detailed models. A classical tool is provided by the so-called source-receptor matrices that are based on the linearization of the function $c = F(E)$ (c are the concentrations, E are the emissions and F is the model viewed as a black box). We refer to [2] for the evaluation of the resulting errors.

An alternative technique is to use highly reduced models. We now present some methods than can be used for the reduction of such high-dimensional systems.

7.2 Slow/fast models and dynamic reduction

A first approach is related to the wide range of timescales that characterizes atmospheric chemistry. Let $\varepsilon \ll 1$ be the ratio of fast timescales to slow timescales. We first assume that the system can be partitioned in a slow part (c_s) and a fast part (c_f). An ideal system is then (after scaling):

$$\varepsilon \frac{\mathrm{d}c_\mathrm{f}}{\mathrm{d}t} = f(c_\mathrm{f}, c_\mathrm{s}) \,, \quad \frac{\mathrm{d}c_\mathrm{s}}{\mathrm{d}t} = g(c_\mathrm{f}, c_\mathrm{s}) \tag{49}$$

Here, f and g stand for the fast and slow dynamics, respectively. A key result given by the singular perturbation theory [139] is that this system may be approximated up to the first order in ε after a fast transient phase by the following differential-algebraic system:

$$0 = f(c_\mathrm{f}, c_\mathrm{s}) \,, \quad \frac{\mathrm{d}c_\mathrm{s}}{\mathrm{d}t} = g(c_\mathrm{f}, c_\mathrm{s}) \tag{50}$$

This is the theoretical framework for the QSSA methods (Section 3). The key point is of course to partition the system under the slow/fast form. Simple criteria can be derived from the production-loss form of chemical kinetics [33]. The use of reduced models (that are no more stiff) is investigated in [56, 65, 75, 76, 87, 126] for gas-phase chemistry and [36] for multiphase chemistry. We also refer to [34] for the numerical solution of the resulting model.

For aerosol modeling, the fast dynamics is related to condensation/evaporation. It appears that the finest aerosols quickly reach equilibrium for mass transfer. This leads to the definition of the so-called *hybrid* methods [17]: the smallest aerosols whose diameter is below a cut-off diameter are supposed to be at equilibrium (they define the fast part), whereas the coarsest aerosols are solved with kinetic mass transfer (they define the slow part). We refer to [28] for a benchmark of different methods and the application to the GDE.

7.3 Building look-up tables

An alternative approach is to replace the most computationally demanding modules of CTMs by look-up tables. There is of course a problem of dimension. If there are n model inputs and s sampling points, the table is of dimension $\mathrm{O}(s^n)$!

A powerful technique is provided by the multivariate expansions, a tool used in many other fields. One example is, for instance, the high-dimensional model representation [101, 119].

Let $c = F(\Psi)$ be the module to be replaced by a look-up table. The Ansatz is that one can write the multivariate exact expansion:

$$F(\Psi) = F_0 + \sum_{i=1}^{i=n} F_i(\Psi_i) + \sum_{1 \le i < j \le n} F_{ij}(\Psi_i, \Psi_j) + \sum_{1 \le i < j < k \le n} F_{ijk}(\Psi_i, \Psi_j, \Psi_k) + \dots + F_{1,2,\dots,n}(\Psi_1, \Psi_2, \dots, \Psi_n) \tag{51}$$

F_i describe the independent action of Ψ_i, F_{ij} is the cooperative action of Ψ_i and Ψ_j, etc. A reduced model is then to use a truncated development (of second order or of third order in practice). The functions F_i are computed such that the truncation error is minimal for a given norm, over the space of functions that satisfy the Ansatz. This procedure can be repeated for each order.

A crucial point is the choice of the error metrics. We refer to [12, 157] for the application to atmospheric chemistry.

7.4 POD/EOF methods

Another method is based on the so-called Proper Orthogonal Decomposition (POD) [6], also known as Empirical Orthogonal Functions in atmospheric modeling (see for instance [44] for an application to the understanding of the surface ozone patterns over the USA). The idea is to map the initial model, let say given by an evolution equation $\frac{dc}{dt} = f(c)$ onto a reduced space, $\text{span}\{\Phi_i\}_{i=1}^{N_r}$:

$$c = \sum_{i=1}^{i=N_r} (c, \Phi_i)\Phi_i \tag{52}$$

where the POD vectors Φ_i are orthonormal with respect to a given scalar product $(.,.)$. The POD basis is defined such that for each $1 \le k \le N_r$, $(\Phi_1 \dots, \Phi_k)$ minimizes the residual J_k for any subspace of dimension k, $\text{span}\{\Phi_i'\}_{i=1}^{k}$:

$$J_k(\Phi_1', \dots, \Phi_k') = \int_0^T \|c(t) - \sum_{i=1}^{i=k} \left(\Phi_i', c(t)\right) \Phi_i'\|^2 \mathrm{d}t \tag{53}$$

along a trajectory $c(t)$. The norm $\|.\|$ is related to the scalar product $(.,.)$. A classical result is that the POD basis can be computed on the basis of the eigenvalues of the correlation matrix $\left[\left(c(t_i), c(t_j)\right)\right]_{ij}$. The choice of the scalar product is once more a challenging point.

In [128], the application to comprehensive chemical mechanism for tropospheric chemistry indicates that there are only 2 or 3 local degrees of freedom for a species like ozone. This result is promising, but the extension to 3D modeling is still an open question.

8 Architecture

These topics illustrate the growing complexity of issues related to the CTMs. To conclude, we would like to emphasize that this provides constraints for the architecture of modeling systems.

The key point is to have highly modular systems rather than all-in-one models because of at least three motivations. First, there exist many available parameterizations for a given process, and the model configuration should be able to take any set of acceptable physical parameterizations. Second, high-level methods such as data assimilation or uncertainties propagation techniques are not related to specific models. At last, one can expect that multimodeling approaches and ensemble forecast are the future of modeling in the field.

An example of a system that meets these requirements (it has been designed in this aim) is the modeling platform POLYPHEMUS (www.enpc.fr/cerea/polyphemus, [78]). POLYPHEMUS is composed of four autonomous levels:

- A C++ Library for parameterizations and preprocessing (ATMODATA)
- A set of drivers for high-level uses (sensitivity study, ensemble, Monte Carlo, data assimilation algorithms, ...)
- A set of numerical models (limited to solvers), for instance, the 3D chemistry-transport-model POLAIR3D [13]
- A postprocessing tool with a focus on model-to-data error statistics (especially those recommended by the US Environmental Protection Agency) and ensemble algorithms, the Python module ATMOPY.

One can expect that similar modeling systems will replace all-in-one models in the near future.

9 Conclusion

We have reviewed in this paper some current issues for air quality modeling and simulation. There are still many challenging topics for applied mathematics ranging from mathematical physics (model formulation for subgrid parameterizations) to numerics, computational physics, etc.

In the near future, the first trend will be to increase the resolution of CTMs. For chemistry, "true" chemical species will replace "model" lumped species (especially for a better representation of organic chemistry). At continental scale, the current model resolution ranges from 10 to 100 km and will go down to 5-10 km. This point is of course related to the improvement of the meso-scale meteorological models and of the emission inventory. The complexity of physical models, especially for aerosols, will increase: a better understanding of secondary organic aerosol implies a better representation of organic species (more variables) [53], a coupling between organic and inorganic thermodynamics (more computational burden) [98] and new processes (surface heterogeneous chemistry).

A second point is related to the couplings and the feedbacks. This concerns, for instance, the coupling between radiative transfer models, CTM and meso-scale models for the radiative behavior of the atmosphere. Another example is provided by detailed multimedia models in the estimation of environmental impacts. This will require the development of new numerical algorithms with less splitting.

One must keep in mind that the final use of such models always needs a drastic reduction of the CPU costs. Finding the appropriate trade-off between highly detailed physical models (the "arms race") and tailored models that can be used in integrated modeling for impact assessment will still remain a key issue. The current situation is that one *often* uses low-level models and *sometimes* detailed models for such applications. Models of intermediate complexity or algorithms for using detailed models are therefore of great interest, all the more that uncertainties have to be taken into account through the computation of PDFs.

Acknowledgements I thank many of my colleagues and my former PhD students for their contributions to some of the referenced works and to my own understanding of air quality modeling and simulation: Vivien Mallet for many topics (CTMs, uncertainties, ensemble forecast), Denis Quélo (inverse modeling), Edouard Debry and Karine Sartelet (aerosol modeling), Rafik Djouad (aqueous-phase chemistry) and Jaouad Boutahar (proxy models).

References

1. Babovsky, H.: On a Monte Carlo scheme for Smoluchowski's coagulation equation. Monte Carlo Methods Appl. **5**(1), 1–18 (1999)
2. Bartnicki, J.: Nonlinear effects in the source-receptor matrices computed with the EMEP Eulerian acid deposition model. Technical Report 4/2000, EMEP/MSC-W (2000)
3. Bennett, A.: Inverse Modeling of the ocean and atmosphere. Cambridge University Press, Cambridge (2002)
4. Bergamaschi, P., Hein, R., Heimann, M., Crutzen, P.: Inverse modeling of the global CO cycle. 1. Inversion of CO mixing ratios. J. Geophys. Res. **105**(D2), 1909–1927 (2000)
5. Bergamaschi, P., Krol, M., Dentener, F., Vermeulen, A., Meinhardt, F., Graul, R., Ramonet, M., Peteres, W., Dlugokencky, E.: Inverse modelling of national and european CH_4 emissions using the atmospheric zoom model TM5. Atmos. Chem. Phys. **5**, 2431–2460 (2005)
6. Berkooz, G., Holmes, P., Lumley, J.: The proper orthogonal decomposition in the analysis of turbulent flows. Annu. Rev. Fluid Mech. 539–575 (1993)
7. Berkvens, P., Botchev, M., Krol, M., Peters, W., Verwer, J.: Solving vertical transport and chemistry in air pollution models. In: Chock, D., Carmichael, G. (eds.): Atmospheric Modeling, Springer, Berlin Heidelberg New York (2002)
8. Berliner, L., Lu, Z., Snyder, C.: Statistical design for adaptive weather observations. J. Atmos. Sci. **56**, 2536–2552 (1999)
9. Binkowski, F., Roselle, S.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component. 1. Model description. J. Geophys. Res. **108**(D6-4183), DOI: 10.1029/2001JD001409 (2003)
10. Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy I : Theories. Quart. J. Roy. Meteorol. Soc. **131**(Part B(610)) (2005a)
11. Bocquet, M.: Reconstruction of an atmospheric tracer source using the principle of maximum entropy II : Applications. Quart. J. Roy. Meteorol. Soc. **131**(Part B(610)) (2005b)
12. Boutahar, J.: Reduction and evaluation of uncertainties in chemistry-transport models. Ph.D. thesis, ENPC. In French (2004)
13. Boutahar, J., Lacour, S., Mallet, V., Quélo, D., Roustan, Y., Sportisse, B.: Development and validation of a fully modular platform for numerical modelling of air pollution: Polair3D. Int. J. Environ. Pollut. **22**(1-2) (2004)
14. Boutahar, J., Sportisse, B.: Reduction methods for atmospheric chemistry. In: Global and Regional Atmospheric Modeling Meeting, pp. 133–139. University of Aveiro, Portugal (2002)
15. Byrne, G., Hindmarsh, A.: Stiff ODE solvers: a review of current and coming attractions. J. Comp. Phys. **70**, 1–62 (1987)
16. Byun, D., Lee, S.: Numerical solution of trace species advection under non-uniform density distribution: experiment with two-dimensional linear flows. In: Chock, D., Carmichael, G. (eds.): Atmospheric Modeling, pp. 109–151. Springer, Berlin Heidelberg New York (2002)
17. Capaldo, K., Pilinis, C., Pandis, S.: A computationally efficient hybrid approach for dynamic gas/aerosol transfer in air quality models. Atmos. Environ. **34**, 3617–3627 (2000)
18. Carmichael, G., et al.: Sensitivity analysis for atmospheric chemistry models via automatic differentiation. Atmos. Environ. **31**(3), 475–489 (1997)
19. Cesa-Bianchi, N., Long, P.M., Warmuth, M.K.: Worst-case quadratic loss bounds for prediction using linear functions and gradient descent. IEEE Trans. Neural Networks **7**(3), 604–619 (1996)
20. Chang, M., Hartley, D., Cardelino, C., Chang, W.-L.: Inverse modeling of biogenic emissions. Geophys. Res. Lett. **23**, 3007 (1996)
21. Chang, M., Hartley, D., Cardelino, C., Haas-Laursen, D., Chang, W.-L.: On using inverse methods for resolving emis-

sions with large spatial inhomogeneities. J. Geophys. Res. **102**(D13), 16023–16036 (1997)

22. Constantinescu, E.M., Sandu, A., Chai, T., Carmichael, G.R.: Assessment of ensemble-based chemical data assimilation in an idealized setting. Atmos. Environ. **41**(1), 18–36 (2007)
23. Dabdub, D., Nguyen, K.: Semi-lagrangian flux scheme for the solution of the aerosol condensation/evaporation equation. Aerosol Sci. Tech. **36**, 407–418 (2002)
24. Daescu, D., Sandu, A., Carmichael, G.: Direct and adjoint sensitivity analysis of chemical kinetics with KPP: II Numerical validation and applications. Atmos. Environ. (2003)
25. Debry, E.: Numerical simulation of an atmospheric aerosol distribution. Ph.D. thesis, ENPC, CEREA. In French (2004)
26. Debry, E., Fahey, K., Sartelet, K., Sportisse, B., Tombette, M.: A new size resolved aerosol model: SIREAM. Atmos. Chem. Phys. Discuss. **6**, 11845–11875 (2006)
27. Debry, E., Sportisse, B.: Numerical simulation of the General Dynamics Equation (GDE) for aerosols with two collocation methods. Appl. Numer. Math. (In press)
28. Debry, E., Sportisse, B.: Reduction of the condensation/evaporation dynamics for atmospheric aerosols: theoretical and numerical investigation of hybrid methods. J. Aerosol Sci. **37**(8), 950–966 (2006b)
29. Debry, E., Sportisse, B.: Solving aerosol coagulation with size-binning methods. Appl. Numer. Math. (In press)
30. Debry, E., Sportisse, B., Jourdain, B.: A stochastic approach for the numerical simulation of the general dynamics equation for aerosols. J. Comp. Phys. **184**, 649:689 (2003)
31. Dhaniyala, S., Wexler, A.: Numerical schemes to model condensation and evaporation of aerosols. Atmos. Environ. **30**(6), 919–927 (1995)
32. Djouad, R., Audiffren, N., Sportisse, B.: Sensitivity analysis using automatic differentiation applied to a multiphase chemical mechanism. Atmos. Environ. **37**(22), 3029–3038 (2003a)
33. Djouad, R., Sportisse, B.: Partitioning techniques for reduction in chemical kinetics. APLA: an automatic partitioning and lumping algorithm. Appl. Numer. Math. **43**(4), 383–398 (2002)
34. Djouad, R., Sportisse, B.: Solving reduced models in Air pollution modelling. Appl. Numer. Math. **44**(1), 49–61 (2003)
35. Djouad, R., Sportisse, B., Audiffren, N.: Numerical simulation of aqueous-phase atmospheric models : Use of a non-autonomous Rosenbrock method. Atmos. Environ. **36**, 873–879 (2002)
36. Djouad, R., Sportisse, B., Audiffren, N.: Reduction of multiphase atmospheric chemistry. J. Atmos. Chem. **46**, 131–157 (2003b)
37. Domilovskii, E., Lushnikov, A., Piskunov, V.: Monte Carlo simulation of coagulation processes. Dokl. Akad. Nauk SSSR, Ser. Phys. Chem. **240**(1), (1978)
38. Eibeck, A., Wagner, W.: Stochastic particle approximations for Smoluchowski's coagulation equation. Technical report, Weierstrass-Institut for Applied Analysis and Stochastics. Preprint No. 585 (2000)
39. Elbern, H., Schmidt, H.: A four dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling. J. Geophys. Res. **104**(18), 583–18598 (1999)
40. Elbern, H., Schmidt, H.: Ozone episode analysis by four dimensional variational chemistry data assimilation. J. Geophys. Res. **106**, 3569–3590 (2001)
41. Elbern, H., Schmidt, H., Ebel, A.: Variational data assimilation for tropospheric chemistry modeling. J. Geophys. Res. **102**, 15967–15985 (1997)
42. Elbern, H., Schmidt, H., Talagrand, O., Ebel, E.: 4D variational data assimilation with an adjoint air quality model for emission analysis. Environ. Model. Softw. **15**, 539–548 (2000)
43. Fife, P.: Mathematical aspects of reacting and diffusing systems, vol. 28. Springer, Berlin Heidelberg New York (1979)
44. Fiore, A., Jacob, D.: Application of empirical orthogonal functions to evaluate ozone simulations with regional and global models. J. Geophys. Res. **108**(D14) (2003)
45. Frohn, L., Christensen, J., Brandt, J.: Development of a high-resolution nested air pollution model. J. Comp. Phys. **179**, 68–94 (2002)
46. Galmarini, S., Bianconi, R., Klug, W., Mikkelsen, T., Addis, R., Andronopoulos, S., Astrup, P., Baklanov, A., Bartniki, J., Bartzis, J.C., et al.: Ensemble dispersion forecasting – Part I: concept, approach and indicators. Atmos. Environ. **38**(28), 4607–4617 (2004)
47. Gelbard, F., Seinfeld, J.: Numerical solution of the dynamic equation for particulate systems. J. Comp. Phys. **28**, 357–375 (1977)
48. Gelbard, F., Seinfeld, J.: The general dynamics equation for aerosols. J. Colloid Interface Sci. **68**(2), 363–382 (1979)
49. Giering, R., Kaminski, T.: Recipes for adjoint code construction. ACM Trans. Math. Software **24**(4), 437–474 (1998)
50. Gong, Cho: A numerical scheme for the integration of the gas-phase chemical rate equations in 3D atmospheric models. Atmos. Environ. **27A**, 2591–2611 (1993)
51. Graf, J., Moussiopoulos, N.: Intercomparison of two models for the dispersion of chemically reacting pollutants. Beitr. Phys. Atmos. **64**(1), 13–25 (1991)
52. Griewank, A., Walther, A.: Revolve: An implementation of checkpoint for the reverse or adjoint mode of computational differentiation. ACM Trans. Math. Software (1999)
53. Griffin, R., Dabdub, D., Seinfeld, J.: Secondary organic aerosol 1. Atmospheric chemical mechanism for production of molecular constituents. J. Geophys. Res. **107**(D17), (2002)
54. Hanna, S.R., Lu, Z., Frey, H.C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., Hansen, D.A.: 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), 891–903 (2001)
55. He, S., et al: Application of ADIFOR for air pollution model sensitivity studies. Environ. Model. Softw. **15**, 475–489 (2000)
56. Heard, A., Pilling, M., Tomlin, A.: Mechanism reduction techniques applied to tropospheric chemistry. Atmos. Environ. **32**(6), 1059–1073 (1998)
57. Hertel, O., Berkowicz, R., Christensen, J.: Test of two numerical schemes for use in atmospheric transport-chemistry models. Atmos. Environ. **27A**(16), 2591–2611 (1993)
58. Higdon, R., Bennett, A.: Stability analysis of operator splitting for large scale ocean modeling. J. Comp. Phys. **123**, 311–329 (1996)
59. Hourdin, F., Issartel, J.-P.: Sub-surface nuclear tests monitoring through the CTBT^{133} Xe network. Geophys. Res. Lett. **27**, 2245–2248 (2000)
60. Hu, Y., Odman, M., Russel, A.: Mass conservation in the CMAQ model. Atmos. Environ. **40**, 1199–1204 (2006)
61. Hundsdorfer, W., Verwer, J.: Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations, Vol. 33 of Springer Series in Computational Mathematics. Springer, Berlin Heidelberg New York (2003)

62. Jacobson, M., Turco, R.: Simulating condensational growth, evaporation and coagulation of aerosols using a combined moving and stationary size grid. Aerosol Sci. Tech. **22**, 73–92 (1995)
63. Jacobson, M., Turco, R., Jensen, E., Toon, O.: Modeling coagulation among particles of different composition and size. Atmos. Environ. **28**(7), 1327–1338 (1994)
64. Jay, L., Sandu, A., Potra, F., Carmichael, G.: Improved QSSA methods for atmospheric chemistry integration. SIAM J. Sci. Comp. (1997)
65. Kalachev, L., Field, R.: Reduction of a model describing ozone oscillations in the troposphere: example of an algorithmic approach to model reduction in atmospheric chemistry. J. Atmos. Chem. **39**, 65–93 (2001)
66. Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge University Press, Cambridge (2003)
67. Karamchandani, P., Santos, L., Sykes, I., Zhang, Y., Tonne, C., Seigneur, C.: Development and application of a state-of-the-science reactive plume model. Environ. Sci. Technol. **34**, 870–880 (2000)
68. Kim, Y., Seinfeld, J.: Simulation of multicomponent aerosol condensation by the moving sectional method. J. Colloid Interface Sci. **135**(1), 185–199 (1990)
69. Krishnamurti, T., Kishtawal, T., LaRow, T., Bachiocchi, D., Zhang, Z., Williford, C., Gadhil, S., Surendran, S.: Improved weather and seasonal climate forecasts from multi-model superensemble. Science **285**, 1548–1550 (1999)
70. Lamb, R.: Note on the application of K-theory to diffusion problems involving nonlinear chemical reactions. Atmos. Environ. **7**, 257–263 (1973)
71. Lanser, L., Verwer, J.: Analysis of operator splitting for advection–diffusion–reaction problems from air pollution modelling. In: Proceedings 2nd Meeting on Numerical methods for differential equations (1998)
72. Larson, V., Golaz, J.-C., Jiang, H., Cotton, W.: Supplying local microphyisics parameterizations with information about subgrid variability: Latin Hypercube sampling. J. Atmos. Sci. **62**, 4010–4026 (2005)
73. Lee, S., Yoon, S., Byun, D.: The effect of mass inconsistency of the meteorological field generated by a common meteorological model on air quality modeling. Atmos. Environ. **38**, 2917–2926 (2004)
74. Louis, J.: A parametric model of vertical eddy fluxes in the atmosphere. Bound.-Lay. Meteorol. **17**, 197–202 (1979)
75. Lovas, T., Mastorakos, E., Goussis, D.: Reduction of the RACM scheme using computational singular perturbation analysis. J. Geophys. Res. **111**(D13302), (2006)
76. Lowe, R., Tomlin, A.: Low-dimensional manifolds and reduced chemical models for tropospheric chemistry simulations. Atmos. Environ. **34**, 2425–2436 (2000)
77. Mallet, V., Sportisse, B.: 3D Chemistry-Transport Model Polair3D: numerical issues, validation and automatic-differentiation strategy. Atmos. Chem. Phys. Discuss. (1), 1371–1392 (2004)
78. Mallet, V., Sportisse, B.: A comprehensive study of ozone sensitivity with respect to emissions over Europe with a chemistry-transport model. J. Geophys. Res. **110**(D22), (2005)
79. Mallet, V., Sportisse, B.: Ensemble-based air quality forecasts: a multi-model approach applied to ozone. J. Geophys. Res. **111**(D18), 18302 (2006a)
80. Mallet, V., Sportisse, B.: Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: an ensemble approach applied to ozone modeling. J. Geophys. Res. **111**(D01302) (2006b)
81. Marchuk, G.: Mathematical models in environmental problems, vol. 16. North Holland, Amsterdam (1986)
82. McRae, G., Goodin, W., Seinfeld, J.: Numerical solution of the atmospheric diffusion equation for chemically reacting flows. J. Comp. Phys. **45**, 1–42 (1982)
83. Meirink, J., Eskes, H., Goede, P.: Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling. Atmos. Chem. Phys. Discuss. **5**, 9405–9445 (2005)
84. Müller, F.: Splitting error estimation for microphysical-multiphase chemical systems in meso-scale air quality models. Atmos. Environ. 5749–5764 (2001)
85. Molemaker, M., Vila-Guerau de Arellano, J.: Control of chemical reactions by convective turbulence in the boundary layer. J. Atmos. Sci. **55**, 568–579 (1998)
86. Mott, D., Oran, E., Van Leer, B.: A Quasi-Steady-State solver for the stiff Ordinary Differential Equations of reaction kinetics. J. Comp. Phys. **164**, 407–428 (2000)
87. Neophytou, M., Goussis, D., Van Loon, M., Mastorakos, E.: Reduced chemical mechanisms for atmospheric pollution using computational singular perturbation analysis. Atmos. Environ. **38**, 3661–3673 (2004)
88. Nguyen, K., Dabdub, D.: Development and analysis of a non-splitting solution for three-dimensional air quality models. Atmos. Environ. **37**, 3741–3748 (2003)
89. Odman, M., Russel, A.: Mass conservative coupling of non-hydrostatic meteorological models with air quality models. In: Gryning, S., Batchvarova, E. (eds.): Air Pollution Modeling and Its Applications, pp. 651–660. Plenum Press, New York (2000)
90. Pagowski, M., Grell, G.A., McKeen, S.A., Dévényi, D., Wilczak, J.M., Bouchet, V., Gong, W., McHenry, J., Peckham, S., McQueen, J., Moffet, R., Tang, Y.: A simple method to improve ensemble-based ozone forecasts. Geophys. Res. Lett. **32**, (2005)
91. Pandis, S., Seinfeld, J.: Sensitivity analysis of a chemical mechanism for aqueous-phase atmospheric chemistry. J. Geophys. Res. **94**(D1), 1105–1126 (1989)
92. Petersen, A., Spee, E., Van Dop, H., Hundsdorfer, W.: An evaluation and intercomparison of four new advection schemes for use in global chemistry models. J. Geophys. Res. **103**, 19253–19270 (1998)
93. Pham, D.: A singular evolutive interpolated Kalman filter for data assimilation in oceanography. Technical Report RT 163, LMC/IMAG (1996)
94. Pham, D., Verron, J., Roubaud, M.C.: A singular evolutive extended Kalman filter for data assimilation in oceanography. J. Marine Syst. **16**, 323–340 (1998)
95. Pilinis, C.: Derivation and numerical solution of the species mass distribution equations for multicomponent particulate systems. Atmos. Environ. **24A**(7), 1923–1928 (1990)
96. Pilinis, C., Capaldo, K., Nenes, A., Pandis, S.: MADM – a new Multi-component aerosol dynamic model. Aerosol Sci. Tech. **32**, 482–502 (2000)
97. Preussner, P., Brand, K.: Application of a semi-implicit Euler method to mass action kinetics. Chem. Eng. Sci. **36**(10), 1633–1641 (1981)
98. Pun, B., Griffin, R., Seigneur, C., Seinfeld, J.: Secondary Organic Aerosol 2. Thermodynamic model for gas/particle partitioning of molecular constituents. J. Geophys. Res. **107**(D17), (2002)
99. Quélo, D., Mallet, V., Sportisse, B.: Inverse modeling of NO_x emissions at regional scale over Northern France. Preliminary investigation of the second-order sensitivity. J. Geophys. Res. **110**(D24310), (2005)

100. Quélo, D., Sportisse, B., Isnard, O.: Data assimilation for short-range dispersion of radionuclides: a case study for second-order sensitivity. J. Environ. Radioact. **84**, 393–408 (2005)
101. Rabitz, H., Alis, O.: General foundations of high-dimensional model representations. J. Math. Chem. **25**, 197–223 (1999)
102. Ramabhadran, T., Peterson, T., Seinfeld, J.: Dynamics of Aerosol Coagulation and Condensation. AIChe J. **22**(5), 840–851 (1976)
103. Rodriguez, M., Dabdub, D.: Monte Carlo uncertainty and sensitivity analysis of the CACM chemical mechanism. J. Geophys. Res. **108**(D15), (2003)
104. Roustan, Y., Bocquet, M.: Sensitivity analysis for mercury over Europe. J. Geophys. Res. **111**(D14304), (2006)
105. Russell, A., Dennis, R.: NARSTO critical review of photochemical models and modeling. Atmos. Environ. **34**, 2283–2234 (2000)
106. Sandu, A.: Discretizing Aerosol Dynamics with B-Splines. Technical report, Michigan Technological University, Houghton, MI 49931 (2001a)
107. Sandu, A.: Positive numerical integration methods for chemical kinetic systems. J. Comp. Phys. **170**, 1–14 (2001b)
108. Sandu, A.: Targeted Observations for Atmospheric Chemistry and Transport Models. In: International Conference for Computational Science (ICCS-2006) (2006)
109. Sandu, A., Borden, C.: A framework for the numerical treatment of aerosol dynamics. Appl. Numer. Math. **45**, 475–497 (2003)
110. Sandu, A., Daescu, D., Carmichael, G.: Direct and Adjoint sensitivity analysis of chemical kinetics with KPP: I Theory and software tools. Atmos. Environ. (2003)
111. Sandu, A., Daescu, D., Carmichael, G., Chai, T.: Adjoint sensitivity analysis of regional air quality models. J. Comp. Phys. **204**, 222–252 (2005)
112. Sandu, A., Potra, F., Carmichael, G., Damian, V.: Efficient implementation of fully implicit methods for atmospheric chemical kinetics. J. Comp. Phys. **129**, 101–110 (1996)
113. Sandu, A., Verwer, J., Blom, J., Spee, E., Carmichael, G.: Benchmarking stiff ODEs solvers for atmospheric chemistry problems II: Rosenbrock solvers. Atmos. Environ. **31**, 3459–3472 (1997a)
114. Sandu, A., Verwer, J., Van Loon, M., Carmichael, G., Potra, F., Dabdub, D., Seinfeld, J. : Benchmarking stiff ODEs solvers for atmospheric chemistry problems I: implicit versus explicit. Atmos. Environ. **31**, 3151–3166 (1997b)
115. Sartelet, K.N., Hayami, H., Albriet, B., Sportisse, B.: Development and preliminary validation of a modal aerosol model for tropospheric chemistry: MAM. Aerosol Sci. Tech. **40**(2), 118–127 (2005)
116. Segers, A.: Data assimilation in atmospheric chemistry models using Kalman filtering. Ph.D. thesis, Delft University (2002)
117. Seigneur, C., Stephanopoulos, G., Carr Jr, R.: Dynamic sensitivity analysis of chemical reaction systems. Chem. Eng. Sci. **37**(6), 845–853 (1982)
118. Seinfeld, J., Pandis, S.: Atmospheric Chemistry and Physics. Wiley-Interscience, New York (1998)
119. Shorter, J., Ip, P.C.: An efficient chemical kinetics solver using high dimensional model representation. J. Phys. Chem. (1999) **103**, 7192–7198 (1999)
120. Sirkes, Z., Tziperman, E.: Finite difference of adjoint or adjoint of finite difference?. Mon. Weather Rev. **125**, 3373–3378 (1997)
121. Sommeijer, B., Van der Houwen, P., Verwer, J.: On the treatment of time-dependent boundary conditions in splitting methods for parabolic differential equations. Int. J. Num. Met. Eng. **17**, 335–346 (1981)
122. Sportisse, B.: An analysis of operator splitting techniques in the stiff case. J. Comp. Phys. **161**, 140–168 (2000)
123. Sportisse, B.: Box models versus Eulerian models in air pollution modeling. Atmos. Environ. **35**, 173–178 (2001)
124. Sportisse, B.: A review of parameterizations for modeling dry deposition and scavenging of radionuclides. Atmos. Environ. (In press)
125. Sportisse, B., Bencteux, G., Plion, P.: Method of lines versus operator splitting methods for reaction–diffusion systems in air pollution modelling. Environ. Model Softw. **15**, (2000)
126. Sportisse, B., Djouad, R.: Reduction of chemical kinetics in air pollution modelling. J. Comp. Phys. **164**, 354–376 (2000)
127. Sportisse, B., Djouad, R.: Mathematical investigation of mass transfer for atmospheric pollutants into a fixed droplet with aqueous chemistry. J. Geophys. Res. **108**(D2), 4073 (2003)
128. Sportisse, B., Djouad, R.: Use of Proper Orthogonal Decompositions for the reduction of atmospheric chemistry. J. Geophys. Res. (In press)
129. Sportisse, B., Du Bois, L.: Numerical and theoretical investigation of a simplified model for the parameterization of below-cloud scavenging by falling raindrops. Atmos. Environ. **36**, 5719–5727 (2002)
130. Sportisse, B., Quélo, D.: 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 (2002)
131. Sportisse, B., Quélo, D., Mallet, V.: Impact of mass consistency errors for atmospheric dispersion. Atmos. Environ. (In press)
132. Srivastava, R., Mc Rae, D., Odman, M.: Simulation of dispersion of a power plant plume using an adaptive grid algorithm. Atmos. Environ. **35**, 4801–4818 (2001)
133. Stockwell, W., Kirchner, F., Kuhn, M., Seefeld, S.: A new mechanism for regional atmospheric chemistry modeling. J. Geophys. Res. **95**(D10), 16343–16367 (1997)
134. Strang, G.: On the construction and comparison of difference schemes. SIAM J. Numer. Anal. **5**, 506–517 (1968)
135. Stratton, D., Gans, J., Williams, E.: Coagulation algorithms with size binning. J. Comp. Phys. **112**, 364–369 (1994)
136. Sun, P.: A pseudo non-time splitting method in air quality modeling. J. Comp. Phys. **127**, 152–157 (1996)
137. Tarantola, A.: Inverse Problem Theory. Elsevier, Amsterdam (1987)
138. Tatang, M., Pan, W., Prinn, R., McRae, G.: An efficient implementation for parametric uncertainty analysis of numerical geophysical models. J. Geophys. Res. **102**, 21925–21932 (1997)
139. Tikhonov, A., Vasileva, A., Sveshnikov, A.: Differential equations. Springer, Berlin Heidelberg New York (1985)
140. Tomlin, A., Berzins, M., Ware, J., Smith, J., Pilling, M.: On the use of adaptative gridding methods for modelling chemical transport from multi-scales sources. Atmos. Environ. **31**(18), 2945–2959 (1997)
141. Troen, I.B., Mahrt, L.: A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. Bound.-Lay. Meteorol. **37**, 129–148 (1986)
142. Tzivion, S., Reisin, T., Levin, Z.: A numerical solution of the kinetic collection equation using high spectral grid resolution : A proposed reference. J. Comp. Phys. **148**, 527–544 (1999)

143. Van Loon, M.: Numerical methods in smog prediction. Ph.D. thesis, Univ. Amsterdam (1996)
144. Vasileva, A., Butuzov, V., Kalachev, L.: The boundary function method for singular perturbation problems. SIAMs Monographs (1994)
145. Verlaan, M.: Efficient Kalman filtering algorithms for hydrodynamic models. Ph.D. thesis, TU Delft (1998)
146. Verwer, J.: Gauss-Seidel iteration for stiff ODEs from chemical kinetics. SIAM J. Sci. Comp. **15**, 1243–1250 (1994)
147. Verwer, J., Blom, J., Hundsdorfer, W.: An implicit-explicit approach for atmospheric transport-chemistry problems. Appl. Numer. Math. **20**, 191–209 (1996a)
148. Verwer, J., Blom, J., Van Loon, M., Spee, E.: A comparison of stiff ODEs solvers for atmospheric chemistry problems. Atmos. Environ. **30**, 49–58 (1996b)
149. Verwer, J., Hundsdorfer, W., Blom, J.: Numerical time integration for air pollution models. In: Proceedings of the Conference APMS'98 (1998)
150. Verwer, J., Simpson, D.: Explicit methods for stiff ODEs from atmospheric chemistry. Appl. Numer. Math. **18**, 413–430 (1995)
151. Verwer, J., Spee, E., Blom, J., Hundsdorfer, W.: A second order Rosenbrock method applied to photochemical dispersion problem. SIAM J. Sci. Comp. **20**(4), 1456–1480 (1999)
152. Verwer, J., Sportisse, B.: A note on Operator Splitting analysis in the stiff linear case. Technical Report MAS-R9830, CWI (1998)
153. Verwer, J., Van Loon, M.: An evaluation of explicit Pseudo-Steady State Approximation schemes for stiff ODE systems from chemical kinetics. J. Comp. Phys. **113**, 347–352 (1994)
154. Vila-Guerau de Arellano, J., Dosio, A., Vinuesa, J., Holtslag, A., Galmarini, S. : The dispersion of chemically reactive species in the atmospheric boundary layer. Meteorol. Atmos. Phys. **87**(23-38), (2004)
155. Vinuesa, J.-F., Vila-Guerau de Arellano, J.: Fluxes and covariances of reacting scalars in the convective boundary layer. Tellus **B 55**(4), 935–949 (2003)
156. Vinuesa, J.-F., Vila-Guerau de Arellano, J.: Introducing effective reaction rates to account for the inefficient mixing of the convective boundary layer. Atmos. Environ. **39**, 445–461 (2005)
157. Wang, S., Levy II, H., Li, G., Rabitz, H.: Fully equivalent operational models for atmospheric chemical kinetics within global chemistry-transport models. J. Geophys. Res. **104**(D23), 30417–30426 (1999)
158. Wang, Z., Navon, I., Le Dimet, F., Zou, X.: The second order adjoint analysis: theory and applications. Meteorol. Atmos. Phys. **50**, 3–20 (1992)
159. Whitby, E., McMurry, P.: Modal aerosol dynamics modeling. Aerosol Sci. Tech. **27**, 673–688 (1997)
160. Whitby, E., Stratmann, F., Wilck, M.: Merging and remapping modes in a modal aerosol dynamics models: a "Dynamic Mode Manager". J. Aerosol Sci. **33**, 623–645 (2002)
161. Williamson, D.: Time-split versus process-split coupling of parameterizations and dynamical core. J. Climate (2001)
162. Wolke, R., Knoth, O.: An explicit-implicit numerical approach for atmospheric chemistry-transport modeling. Atmos. Environ. **32**, 1785–1797 (1998)
163. Yanenko, N.: The method of fractional steps. Springer, Berlin Heidelberg New York (1971)
164. Yang, B., Pope, S.: An investigation of the accuracy of manifold methods and splitting schemes in the computational implementation of combustion chemistry. Combust. Flame **112**, 16–32 (1998)
165. Young, T., Boris, J.: A numerical technique for solving stiff ODE associated with the chemical kinetics of reaction flow problems. J. Phys. Chem. **81**, 2424 (1977)
166. Zhang, Y., Bischof, C., Easter, R., Wu, P.: Sensitivity analysis of a mixed-phase chemical mechanism using automatic differentiation. J. Geophys. Res. **103**(D15), 18953–18979 (1998a)
167. Zhang, Y., Bischof, C., Easter, R., Wu, P.: Sensitivity analysis of photochemical indicators for O_3 chemistry using automatic differentiation. J. Atmos. Chem. **51**, 1–41 (2005)
168. Zhang, Y., Seigneur, C., Seinfeld, J., Jacobson, M., Binkowski, F.: Simulation of Aerosol dynamic: A comparative review of algorithms used in air quality models. Aerosol Sci. Tech. **31**, 487–514 (1999)
169. Zhang, Y., Seinfeld, J., Jacobson, M., Clegg, S., Binkowski, F.: A comparative review of inorganic aerosol thermodynamic equilibrium modules: Similarities, differences and their likely causes. Atmos. Environ. **34**, 117–137 (2000)
170. Zlatev, Z.: Computer Treatment of Large Air Pollution Models. Kluwer Academic Publishers, Dordrecht (1995)