

Mathematical treatment of environmental models

Zahari Zlatev, István Faragó and Ágnes Havasi

Abstract Large-scale environmental models can successfully be used in different important for the modern society studies as, for example, in the investigation of the influence of the future climatic changes on pollution levels in different countries. Such models are normally described mathematically by non-linear systems of partial differential equations, which are defined on very large spatial domains and have to be solved numerically on very long time intervals. Moreover, very often many different scenarios have also to be developed and used in the investigations. Therefore, both the storage requirements and the computational work are enormous. The great difficulties can be overcome only if the following four tasks are successfully resolved: (a) fast and sufficiently accurate numerical methods are to be selected, (b) reliable and efficient splitting procedures are to be applied, (c) the cache memories of the available computers are to be efficiently exploited and (d) the codes are to be parallelized.

Zahari Zlatev

Department of Environmental Science, Aarhus University, Roskilde, Denmark, e-mail: zz@dmu.dk

István Faragó

Department of Applied Analysis and Computational Mathematics and HAS, ELTE Research Group “Numerical Analysis and Large Networks”, Eötvös Loránd University, Budapest, Hungary e-mail: faragois@cs.elte.hu

Ágnes Havasi

Department of Applied Analysis and Computational Mathematics and HAS, ELTE Research Group “Numerical Analysis and Large Networks”, Eötvös Loránd University, Budapest, Hungary e-mail: havasia@cs.elte.hu

1 Description of a large scale environmental model

For the sake of simplicity we shall restrict ourselves on the area of long-range transport of air pollution and to a particular model (UNI-DEM, the Unified Danish Eulerian Model, [9]), but most of the results can easily be extended to other environmental models. UNI-DEM is described mathematically by the following system of partial differential equations (PDEs):

$$\begin{aligned}
 \frac{\partial c_i}{\partial t} = & -u \frac{\partial c_i}{\partial x} - v \frac{\partial c_i}{\partial y} && \text{- horizontal transport} && (1) \\
 & + \frac{\partial}{\partial x} \left(K_x \frac{\partial c_i}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial c_i}{\partial y} \right) && \text{- horizontal diffusion} \\
 & + Q_i(t, x, y, z, c_1, c_2, \dots, c_q) + E_i(t, x, y, z) && \text{- chemical reactions + emission} \\
 & + (k_{1i} + k_{2i})c_i && \text{- dry and wet depositions} \\
 & - w \frac{\partial c_i}{\partial z} + \frac{\partial}{\partial z} \left(K_z \frac{\partial c_i}{\partial z} \right), && \text{- vertical transport} \\
 & i = 1, 2, \dots, q && \text{- number of equations (chemical species)}
 \end{aligned}$$

The different quantities involved in (1) are briefly described below:

- $c_i = c_i(t, x, y, z)$ is the concentration of the chemical species i at point (x, y, z) of the space domain and at time t of the time-interval,
- $u = u(t, x, y, z)$, $v(t, x, y, z)$ and $w = w(t, x, y, z)$ are wind velocities (along the Ox , Oy and Oz directions, respectively) at the spatial point (x, y, z) and time t ,
- $K_x = K_x(t, x, y, z)$, $K_y = K_y(t, x, y, z)$ and $K_z = K_z(t, x, y, z)$ are diffusivity coefficients at the spatial point (x, y, z) and time t (it is often assumed that K_x and K_y are non-negative constants, while the calculation of K_z is normally rather complicated),
- $k_{1i} = k_{1i}(t, x, y, z)$ and $k_{2i} = k_{2i}(t, x, y, z)$ are deposition coefficients (dry and wet deposition respectively) of chemical species i at the spatial point (x, y, z) and time t of the time-interval. It should be mentioned here that for some of the species these coefficients are non-negative constants. The wet deposition coefficients k_{2i} are equal to zero when it is not raining.
- $E_i(t, x, y, z)$ is emission source for chemical species i at the spatial point (x, y, z) and time t of the time-interval.

2 Splitting the model

The mathematical model defined by (1) is normally split (see [9]) into the following three sub-models:

$$\frac{\partial c_i^{(1)}}{\partial t} = -w \frac{\partial c_i^{(1)}}{\partial z} + \frac{\partial}{\partial z} \left(K_z \frac{\partial c_i^{(1)}}{\partial z} \right), \quad (2)$$

$$\frac{\partial c_i^{(2)}}{\partial t} = -u \frac{\partial c_i^{(2)}}{\partial x} - v \frac{\partial c_i^{(2)}}{\partial y} + \frac{\partial}{\partial x} \left(K_x \frac{\partial c_i^{(2)}}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_y \frac{\partial c_i^{(2)}}{\partial y} \right), \quad (3)$$

$$\frac{\partial c_i^{(3)}}{\partial t} = -Q_i(t, x, y, z, c_1^{(3)}, c_2^{(3)}, \dots, c_q^{(3)}) + E_i(t, x, y, z) + (k_{1i} + k_{2i})c_i^{(3)}, \quad (4)$$

The first of these three sub-models describes the vertical exchange. The second sub-model describes the combination of the horizontal transport (the advection) and the horizontal diffusion. The last sub-model describes the chemical reactions together with the emission sources and the deposition terms.

Note that the three sub-models are fully defined by (2)-(4), but the splitting procedure is not. It will be completely determined only when it is explained how these sub-models are combined. The simple sequential splitting procedure is applied in UNI-DEM. It is obtained in the following way. Assume that the space domain is discretized by using a grid with $N_x \times N_y \times N_z$ grid-points, where N_x , N_y and N_z are the numbers of the grid-points along the grid-lines parallel to the Ox , Oy and Oz axes. Assume further that the number of chemical species involved in the model is $N_s = q$. Finally, assume that approximate values of the concentrations (for all species and at all spatial grid-points) have been found for some $t = t_n$. These values can be considered as components of a vector-function $c(t_n, x_i, y_j, z_k) \in \mathbb{R}^{N_x \times N_y \times N_z \times N_s}$. The next time-step, time-step $n + 1$ (at which approximations of the concentrations are found at $t_{n+1} = t_n + \Delta t$ where Δt is some increment), can be performed by solving successively the three sub-models. The values of $c(t_n, x_i, y_j, z_k)$ are used as an initial condition in the solution of (2). The solution of (2) is used as an initial condition of (3). Finally, the solution of (3) is used as an initial condition of (4). The solution of (4) is accepted as an approximation to $c(t_{n+1}, x_i, y_j, z_k)$. In this way, everything is prepared to start the calculations in the next time-step, step $n + 2$.

The major advantage of any splitting procedure based on the above three sub-models is due to the fact that no extra boundary conditions are needed when (2)–(4) are used. This is true not only for the sequential splitting procedure sketched above, but also for any other splitting procedure based on the sub-models defined by (2)–(4).

3 Choice of numerical methods

Assume that the spatial derivatives are discretized by some numerical algorithm (it must be mentioned here that different numerical algorithms can be applied in the different sub-models and this is one of the big advantages of using splitting techniques: for each sub-model one can select the most suitable algorithm). Then the three systems of PDEs represented by (2)–(4) will be transformed into three systems of ODEs (ordinary differential equations):

$$\frac{dg^{(1)}}{dt} = f^{(1)}(t, g^{(1)}), \quad \frac{dg^{(2)}}{dt} = f^{(2)}(t, g^{(2)}), \quad \frac{dg^{(3)}}{dt} = f^{(3)}(t, g^{(3)}). \quad (5)$$

The components of functions $g^{(m)}(t) \in \mathbb{R}^{N_x \times N_y \times N_z \times N_s}$, $m = 1, 2, 3$ are approximations at time t of the concentrations at all spatial grid-points and for all species. The components of functions $f^{(m)}(t) \in \mathbb{R}^{N_x \times N_y \times N_z \times N_s}$, $m = 1, 2, 3$ depend both on quantities involved in the right-hand-side of (1) and on the particular numerical algorithms that are used in the discretization of the spatial derivatives.

A simple linear finite element method is used to discretize the spatial derivatives in (2) and (3). The spatial derivatives can also be discretized by using other numerical methods as, for example, a pseudo-spectral discretization, a semi-Lagrangian discretization (which can be used only to discretize the first-order derivatives, i.e., the advection part should not be combined with the diffusion part when this method is to be applied) and methods producing non-negative values of the concentrations.

The first system of ODEs in (5) can be solved by using many classical time-integration methods. The well-known θ -method is currently used in UNI-DEM.

Predictor-corrector (PC) methods with several different correctors, which are fully discussed in [7], are used in the solution of the second ODE system in (5). The correctors are carefully chosen so that the stability properties of the method can be enhanced. If the code judges the time-stepsize to be too large for the currently used PC method (and may lead to unstable computations), then it switches to a more stable (but also more expensive, because more corrector formulae are used in order to obtain better stability) PC scheme. On the other hand, if the code judges that the stepsize is too small for the currently used PC method, then it switches to a not so stable but more accurate PC scheme (which is using less corrector formulae and, therefore, is less expensive). In this way the code is trying both to keep the same stepsize and to optimize the performance. More details about this strategy can be found in [7].

The solution of the third system in (5) is much more complicated, because this system is both time-consuming and very stiff. Often the QSSA (Quasi-Steady-State-Approximation) method is used in this part of the model. It is simple and relatively stable but not very accurate (therefore it has to be run with a small time-stepsize). An improved QSSA method was implemented in UNI-DEM. The classical numerical methods for stiff ODE systems (such as the Backward Euler Method, the Trapezoidal Rule and Runge-Kutta algorithms) lead to the solution of non-linear systems of algebraic equations and, therefore, they are normally more expensive. On the other hand, these methods can be incorporated with an error control and perhaps with larger time-steps. Partitioning can also be used. Some convergence problems related to the implementation of partitioning have been studied in [8]. More details about the numerical algorithms can be found in [9].

4 Applying parallelization

Another great advantage of using splitting is the appearance of many natural parallel tasks. It is easy to see that (a) the first system in (5) contains $N_x \times N_y \times N_s$ independent tasks (for each chemical compound, each system along a vertical grid-line

can be treated independently), (b) the second system in (5) contains $N_x \times N_y \times N_z$ independent tasks (the chemical compounds at each grid-point can be treated independently of the chemical compounds at the other grid-points) and (c) the third system in (5) contains $N_z \times N_s$ independent tasks (for each chemical compound the system along a horizontal grid-plane can be treated independently). These parallel tasks, which appear in a natural way when any splitting based on (2)–(4) is applied, were efficiently exploited during the parallelization process. Furthermore, standard parallel tools, OpenMP and MPI, have been extensively used. Much more details can be found in [1] and [9].

5 Applications

UNI-DEM has been used in many different studies (many of them are reported in [9]). Investigations of the influence of the climate changes on pollution levels in Europe [6] and Hungary with its surroundings [10] have recently been carried out.

6 Conclusions

Assume that $N_x = N_y = 480, N_z = 10, N_s = 35$ are used (this was the case in [6] and [10]). Then the number of equations is 80640000 and 213120 time-steps are needed to perform calculations with meteorological and emission data covering a whole year. Moreover, calculations over a long time-period (sixteen years) were needed in [6] and [10]. It is clear that it was possible to resolve the enormous tasks only if (a) efficient splitting procedures are used, (b) suitable numerical methods are selected for each sub-model and (c) parallel computations are applied. It should nevertheless be pointed out that further improvements in connection with the tasks related to (a)-(c) are highly desirable.

Much more details about the mathematical treatment of large environmental models can be found in [2]. More precisely the splitting techniques are treated in [3, 4], the organization of parallel computations described in [11] and the handling of the most difficult part, the sub-model containing the chemical reactions is discussed in [5]. Different applications of environmental models are also reported in [2].

References

1. Alexandrov, V., Owczarz, W., Thomsen, P. G., Zlatev, Z.: Parallel runs of large air pollution models on a grid of SUN computers. *Mathematics and Computers in Simulations*, vol. 65, pp. 557–577 (2004)

2. Faragó, I., Havasi, Á., Zlatev, Z. (eds.): *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*. Bentham, 2013.
3. Faragó, I., Havasi, Á., Zlatev, Z.: Implementation of splitting procedures. In: Faragó, I., Havasi, Á., Zlatev, Z. (eds.): *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, pp. 79–125. Bentham (2013).
4. Faragó, I., Havasi, Á., Zlatev, Z.: Application of splitting in an air pollution model. In: Faragó, I., Havasi, Á., Zlatev, Z. (eds.): *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, pp. 126–165. Bentham (2013).
5. Faragó, I., Havasi, Á., Zlatev, Z.: Treatment of the chemical reactions in an air pollution model. In: Faragó, I., Havasi, Á., Zlatev, Z. (eds.): *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, pp. 54–78. Bentham (2013).
6. Zlatev, Z.: Impact of future climate changes on high ozone levels in European suburban areas. *Climatic Change*, vol. 101, pp. 447–483 (2010)
7. Zlatev, Z.: Application of predictor-corrector schemes with several correctors in solving air pollution problems. *BIT*, vol. 24, pp. 700–715 (1984)
8. Zlatev, Z.: Partitioning ODE systems with an application to air pollution models. *Computers and Mathematics with Applications*, vol. 42, pp. 817–832 (2001)
9. Zlatev, Z., Dimov, I.: *Computational and numerical challenges in environmental modelling*. Elsevier, Amsterdam (2006)
10. Zlatev, Z., Havasi, Á., Faragó, I.: Influence of climatic changes on pollution levels in Hungary and its surrounding countries. *Atmosphere*, vol. 2, pp. 201–221 (2011)
11. Zlatev, Z., Georgiev, K., Dimov, I.: Parallel computations in a large-scale air pollution model. In: Faragó, I., Havasi, Á., Zlatev, Z. (eds.): *Advanced Numerical Methods for Complex Environmental Models: Needs and Availability*, pp. 166–201, Bentham (2013).