THE AIR QUALITY MODELING SYSTEM POLYPHEMUS

Vivien Mallet[1,*], Denis Quélo[2], Bruno Sportisse[3], Meryem Ahmed de Biasi[1], Édouard Debry[4], Irène Korsakissok[4], Lin Wu[1], Yelva Roustan[4], Karine Sartelet[4], Marilyne Tombette[4] and Hadjira Foudhil[4]

[1] INRIA, CLIME project (common project ENPC - INRIA) [France]

[2] IRSN [France]

[3] ENPC, laboratory CEREA (joint research laboratory ENPC - EDF R&D), CLIME project [France]

[4] ENPC, laboratory CEREA (joint research laboratory ENPC - EDF R&D) [France]

INTRODUCTION

Polyphemus [Mallet et al., 2007] is an air quality modeling platform which aims at covering the scope and the abilities of modern air quality systems. With two Gaussian models and two Eulerian models, it handles applications from local scale to continental scale for multiple pollutants: passive tracers, radionuclides, photochemical pollutants, aerosols, and (in development) heavy metals and persistent organic pollutants.

The system is designed to deal with model coupling, ensemble forecast, data assimilation and other advanced simulation methods. The platform-oriented design (with multiple models) and the availability of alternative parameterizations enable to generate ensemble forecasts. On top of the model (and independently of the models), the advanced simulation methods are implemented as drivers in which a model is a black box.

The system is developed at École Nationale des Ponts et Chaussées (one of the top French engineering schools, member of the Paris Institute of Technology), INRIA (French National Institute for Research in Computer Science and Control) and Électricité de France R&D. It is supported by IRSN (Institut de Radioprotection et de Sûreté Nucléaire) and INERIS (French National Institute for Industrial Environment and Risks) which are the main French governmental institutes for atmospheric risk management. It is an open source system distributed at `http://cerea.enpc.fr/polyphemus/`.

1 DESIGN: TECHNICAL DESCRIPTION

1.1 Structure

Polyphemus is built around four bases:

1. data management, data processing facilities (input/output operations, coordinate transformations, interpolations, ...);
2. physical parameterizations (cloud attenuation, turbulence closure, parameterizations for aerosols dynamics, ...);
3. numerical solvers for advection, diffusion and chemistry (including multiphase chemistry);
4. high-level methods (chiefly data assimilation, ensemble forecast and model coupling), in which a chemistry-transport model is simply viewed as a function.

The data management and the physical parameterizations are gathered in dedicated libraries, called SeldonData and AtmoData (the latter being dedicated to atmospheric-related operations). They are the prominent components in the preprocessing steps. In fact, a set of programs process the raw input data (meteorological data, land use data, chemical data, ...), primarily through interpolation, and make calls to physical parameterizations. As far as possible, all fields are computed in this preprocessing stage. This means that the physical formulation in use for a simulation is mostly defined in preprocessing steps.

After preprocessing, the numerical model only performs the time integration of the advection-diffusion-reaction equation. Hence the numerical model includes very few physical calculations.

The numerical model is embedded in a driver. A driver controls the model and interacts with it. The model is seen as a black box with an interface to manage it. For instance, the most simple driver performs a forward simulation with basic calls to the model (initialization and time integration over one time step). Complex drivers are implemented for data assimilation (where, for instance, the driver computes the analysis), for model coupling (the driver then manages two models) and other advanced simulation methods.

The last stage is postprocessing, with visualization, statistical analysis, comparison with measurements, ... It is mostly handled by the Polyphemus library AtmoPy.

The overall structure is shown in Figure 1.

1.2 Technical Choices

Polyphemus relies on three computer languages: C++, Python and Fortran 77.

C++ is the main language. It brings advanced object-oriented abilities. These abilities are needed in order to implement convenient data management

[*] *Corresponding author:* Vivien.Mallet@inria.fr

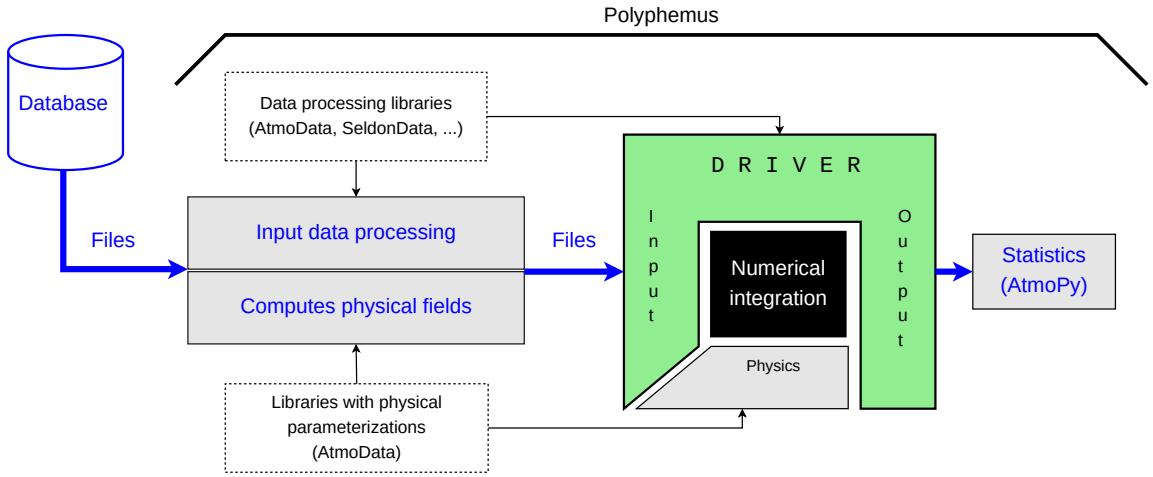

Figure 1: Polyphemus overall work flow. Data is first processed with dedicated C++ libraries (mainly SeldonData and AtmoData). The library AtmoData [Mallet and Sportisse, 2005b; Njomgang et al., 2005] provides in addition most physical parameterizations needed in the preprocessing step. The preprocessing steps output most coefficients of the reactive-transport equation. The numerical model (e.g., an Eulerian chemistry-transport model) integrates in time this equation, and a driver manages the numerical model (e.g., to perform data assimilation). For postprocessing (statistics, visualization, ...), Python scripts and the Python library AtmoPy are used.

and to deal with complex models. The models are C++ objects, that is, black boxes with a clear interface to interact with them and to control them. The drivers therefore manage C++ objects. They are implemented independently of the models; they only require that a model has the right interface, whatever the underlying processes and numerical schemes may be.

Fortran 77 is used for historical reasons and for automatic differentiation (generation of tangent linear model and adjoint model). Python, which is another powerful object-oriented language, covers the other needs, that is, process management and postprocessing (statistics, visualization).

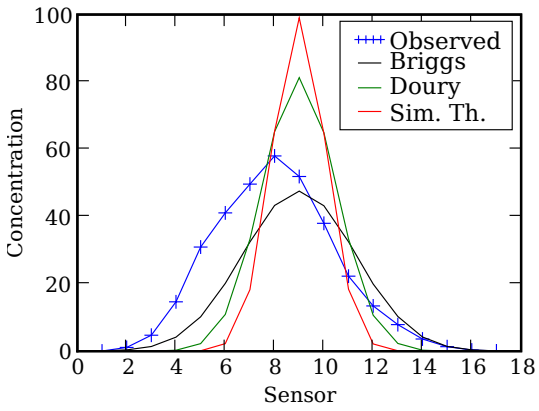

Figure 2: Simulated concentrations and measurements at 100 m from the source and during one Prairie Grass experiment. The simulated concentrations are computed using Briggs formulae (based on Pasquill classes diagnosed with Turner [1969]), Doury formula [Doury, 1976] and a parameterization derived from similarity theory.

2 Contents and Illustrations

2.1 The Models

Polyphemus includes several models:

- a Gaussian plume model;
- a Gaussian puff model;
- an Eulerian chemistry-transport model called Polair3D [Boutahar et al., 2004], relevant up to continental scale;
- an Eulerian chemistry-transport model called Castor, also relevant up to continental scale, clone written in C++ of the gaseous version of Chimere [Schmidt et al., 2001].

The first two models are applied to dispersion at local scale, possibly with radioactive or biological decay. Both models have a gaseous and an aerosol version. They include several alternative physical parameterizations for dispersion coefficients (Figure 2), dry deposition and wet deposition. The Gaussian puff model is also used in a plume-in-grid model (with any of the two Eulerian models).

The Eulerian model Polair3D is derived in three versions: a passive version, a version with gas-phase chemistry (most of the time, radionuclides or

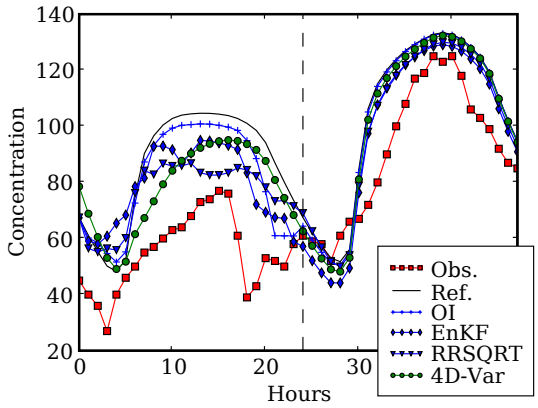

Figure 3: Time evolution of ozone concentrations at EMEP station Montandon, for the reference simulation (without assimilation), the simulation with assimilation (optimal interpolation, OI; ensemble Kalman filter, EnKF; reduced-rank square root Kalman filter, RRSQRT; 4D-Var). The dotted vertical lines delimits the assimilation period.

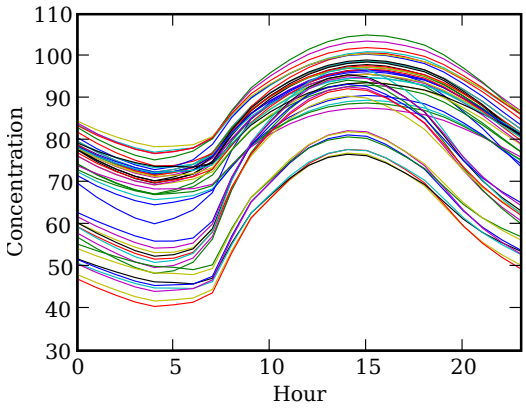

Figure 4: Ozone daily profiles of 48 models built in Polyphemus. The concentrations are in $\mu g\,m^{-3}$ and are averaged over Europe (at ground level) and over summer 2001.

photochemistry) and a version with aerosol dynamics. It may be used with RACM [Stockwell et al., 1997] or RADM 2 [Stockwell et al., 1990] for photochemistry. The aerosol module SIREAM [Debry et al., 2007] is plugged into the aerosol version. Polair3D was evaluated in several studies [e.g., Quélo et al., 2007; Mallet and Sportisse, 2004; Sartelet et al., 2007]. It may be differentiated, essentially with automatic differentiation, to generate a tangent linear model and an adjoint model [Mallet and Sportisse, 2004, 2005a].

2.2 Data Assimilation

Four data assimilation algorithms are available in Polyphemus: optimal interpolation, ensemble Kalman filter [Evensen, 1994], reduced-rank square root Kalman filter [Heemink et al., 2001] and 4D-Var [Le Dimet and Talagrand, 1986]. The Eulerian models Castor and Polair3D have a sufficient interface for sequential algorithms (optimal interpolation and Kalman filters). Only Polair3D may be used in combination with the 4D-Var driver since the latter requires an adjoint model.

A comparison of the four assimilation methods is under progress (Figure 3).

2.3 Ensemble Forecast

At preprocessing stage, most physical fields are computed and they essentially define the physics of a simulation. The model itself only performs the time integration. Polyphemus enables to use alternative parameterizations and alternative data sources in the preprocessing steps. Hence an ensemble of simulations with many different physical choices may be built. This generates a multimodel ensemble [Mallet and Sportisse, 2006b] (Figure 4).

Other ensemble capabilities are available, chiefly Monte Carlo simulations (with a dedicated driver) and model aggregation (“superensembles”, machine learning, . . .) [Mallet and Sportisse, 2006a].

Conclusion

The air quality modeling system Polyphemus is a platform with several chemistry-transport models, with the ability to manage data from several sources, with alternative physical parameterizations for many fields, with several advanced simulation methods (data assimilation, ensemble forecast, model coupling). The multiple choices and methods in the system justify its name: the roots of Polyphemus, in Ancient Greek, mean “multiple speeches”.

Polyphemus structure is open, with the ability to host new models, new physical parameterizations, new advanced simulation methods, . . . Its technical bases enable a strong flexiblity, especially in the management of models. Polyphemus is used in research, but also for in operational forecasts. It is an open source system distributed at `http://cerea.enpc.fr/polyphemus/`. Further details may be found in Mallet et al. [2007] and on Polyphemus website.

References

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

Debry, E., Fahey, K., Sartelet, K., Sportisse, B.,

and Tombette, M. (2007). Technical Note: A new SIze REsolved Aerosol Model (SIREAM). *Atmos. Chem. Phys.*, 7:1,537–1,547.

Doury, A. (1976). Une méthode de calcul pratique et générale pour la prévision numérique des pollutions véhiculées par l'atmosphère. Technical Report 4280, CEA, France.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5):10,143–10,162.

Heemink, A. W., Verlaan, M., and Segers, A. J. (2001). Variance reduced ensemble Kalman filtering. *Mon. Wea. Rev.*, 129:1,718–1,728.

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

Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, É., Korsakissok, I., Wu, L., Roustan, Y., Sartelet, K., Tombette, M., and Foudhil, H. (2007). Technical Note: The air quality modeling system Polyphemus. *Atmos. Chem. Phys. Discuss.*, 7(3):6,459–6,486.

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

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

Mallet, V. and Sportisse, B. (2005b). Data processing and parameterizations in atmospheric chemistry and physics: the AtmoData library. Technical Report 2005-12, CEREA.

Mallet, V. and Sportisse, B. (2006a). Ensemble-based air quality forecasts: A multimodel approach applied to ozone. *J. Geophys. Res.*, 111.

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

Njomgang, H., Mallet, V., and Musson-Genon, L. (2005). AtmoData scientific documentation. Technical Report 2005-10, CEREA.

Quélo, D., Krysta, M., Bocquet, M., Isnard, O., Minier, Y., and Sportisse, B. (2007). Validation of the Polyphemus platform: the ETEX, Chernobyl and Algeciras cases. *Atmos. Env.* In press.

Sartelet, K. N., Debry, E., Fahey, K., Roustan, Y., Tombette, M., and Sportisse, B. (2007). Simulation of aerosols and related species over Europe with the Polyphemus system. Part I: model-to-data comparison for 2001. *Atmos. Env.* Accepted for publication.

Schmidt, H., Derognat, C., Vautard, R., and Beekmann, M. (2001). A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. *Atmos. Env.*, 35:6,277–6,297.

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

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X. (1990). The second generation regional acid deposition model chemical mechanism for regional air quality modeling. *J. Geophys. Res.*, 95(D10):16,343–16,367.

Turner, D. B. (1969). Workbook of atmospheric diffusion estimates. Technical Report 999-AP-26, US EPA, Washington, DC.